THE UNIVERSITY
of EDINBURGH

Dynamics of Numerical Stochastic Perturbation Theory

Marco Garofalo

Doctor of Philosophy
The University of Edinburgh
January 4, 2018

Abstract

Numerical Stochastic Perturbation theory is a powerful tool for estimating high-order perturbative expansions in lattice quantum field theory. The standard algorithm based on the Langevin equation, however, suffers from several limitations which in practice restrict the potential of this technique: first of all it is not exact, a sequence of simulations with finer and finer discretization of the relevant equations have to be performed in order to extrapolate away the systematic errors in the results; and, secondly, the numerical simulations suffer from critical slowing down as the continuum limit of the theory is approached. In this thesis I investigate some alternative methods which improve upon the standard approach. In particular, I present a formulation of Numerical Stochastic Perturbation theory based on the Generalised Hybrid Molecular Dynamics algorithm and a study of the recently proposed Instantaneous Stochastic Perturbation Theory. The viability of these methods is investigated in φ^4 theory.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

Parts of this work have been published in [1] and [2].

(*Marco Garofalo, January 4, 2018*)

Acknowledgements

I am grateful to my supervisor Prof. Anthony Kennedy for the support he has provided throughout my time as his student. His encouragement, knowledge and patient guidance made this work possible.

I want to express my gratitude to Dr. Mattia Dalla Brida who essentially took on the role of a joint supervisor for me. His assistance and expertise have been indispensable throughout all stages of the work.

I thank all my friends and colleagues for all the pleasant times that we have shared. I do not name anyone because I cannot name them all.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Lattice field theory is a tool for understanding the nonperturbative regime of theories as the QCD, but it needs powerful computers to cope with the computational demands. Lattice perturbation theory (LPT) has played an important role in the growth of lattice Field theory (see, e.g., [4, 5] for an introduction). There are many ways in which lattice perturbation theory is useful. It can serve as a direct cross-check of computations carried out in the weak-coupling regime, it can be used in non perturbative calculations to estimate renormalization parameters, it provides a method for computing the matching of physical renormalization schemes employed on the lattice and schemes commonly used in continuum perturbative calculations, such as the $\overline{\mathrm{MS}}$-scheme of dimensional regularization. In addition, LPT gives insight into lattice artefacts of the theory, allowing for both the perturbative determination of Symanzik improvement coefficients and, more generally, of the lattice artefacts in observables of interest.

LPT is technically much more involved than its continuum counterpart because of the complicated form of its vertices and propagators, and usually requires numerical evaluation even for simple diagrams. This is especially evident for some elaborate lattice discretizations or set-up such as, for example the Schrödinger functional [6, 7]. An additional complication arises in the case of

gauge theories, the definition of the gauge-invariant integration measure on the lattice generates new vertices at every order of perturbation theory. This makes the number of diagrams grow very rapidly with the perturbative order, leaving only low-order results accessible to standard techniques.

Numerical stochastic perturbation theory (NSPT) was proposed in [8, 9] (see [10] for a detailed review, and [11–13] for recent developments) in order to circumvent the difficulties of LPT, and thus enable high-order perturbative computations. The basic idea of NSPT is the numerical integration of a discrete version of the equations of stochastic perturbation theory [14] (see [15]). More precisely, starting from the Langevin equation the stochastic field is expanded as a power series in the couplings of the theory and the resulting equations are solved order by order in these couplings. NSPT can be highly automated and can be applied to complicated observables with no additional difficulty. The cost of NSPT scales mildly with the perturbative order, no Feynman diagrams need to be identified or computed, but rather a system of stochastic equations (SDE) is integrated numerically using Monte Carlo techniques. Often NSPT allows high-order perturbative determinations of observables that would not be feasible with other methods.

However NSPT comes with two main disadvantage: firstly, the results at finite lattice spacing unavoidably come with statistical uncertainties due to their Monte Carlo estimation. In particular, the numerical simulations suffer from critical slowing down as the continuum limit of the theory is approached, this significantly increases the computational effort necessary to extract continuum results from NSPT. Secondly, this class of algorithms is not exact: therefore a sequence of simulations with finer and finer discretization of the relevant equations must be performed in order to extrapolate away systematic errors in the results. It is thus difficult to obtain precise results close to the continuum limit for which both systematic and statistical errors are under control. Without continuum extrapolation these methods only provide lattice estimates for perturbative quantities, which in practice may be of limited use.

In the non-perturbative lattice field theory simulations the class of methods

known as Generalized Hybrid Molecular Dynamics (GHMD) algorithms has proven to be superior to Langevin algorithms in this respect; in fact the latter is a special case of the former. It is thus legitimate to check if in the context of NSPT the GHMD algorithms are still superior to the Langevin algorithms.

From a different perspective Martin Lüscher recently introduced a new form of NSPT, namely Instantaneous Stochastic Perturbation Theory (ISPT) [16]. In this work, he discussed how the above limitations can in principle be eliminated completely by formulating NSPT in terms of a certain class of trivializing fields. This method lies somewhere between Langevin NSPT and more conventional diagrammatic perturbation theory.

The aim of this work is to compare the standard NSPT formulation, ISPT, and NSPT based upon GHMD algorithms. Specifically, we will focus on two GHMD algorithms, namely the Hybrid Molecular Dynamics (HMD) algorithm and Kramers algorithm and their NSPT implementations named HSPT and KSPT respectively.

In this work the different NSPT formulations are investigated in the context of φ^4 theory. φ^4 is one of the simplest field theories with interaction terms however, the algorithms are new and there are no data available anywhere. It is sensible to study new algorithms in a simple theory first where good statistics can be achieved relatively cheaply. Also, φ^4 theory is expected to demonstrate all of the important properties of the algorithms. While the exact details will depend on the theory and lattice parameters considered, the continuum scaling of the statistical errors is expected to be the same for different theories.

The structure of the thesis is the following: in §2 we define φ^4 theory on the lattice, the observables used in the numerical investigation, and the renormalization scheme; in §3 we introduce stochastic quantization and we describe the detail of the numerical implementation; in §4 we introduce first the HMD and KSPT, and then the general case of GHMD; in §5 we review ISPT, paying attention to its numerical implementation; in §6 we present results of the numerical investigation of the different methods, followed by our conclusions.

My contribution to this investigation are in: the calculation of the renormalized operator used to compare the different algorithms in LPT up to two loops §2.3 and the argument that it has a well defined continuum limit at all order in perturbation theory §2.2; the alternative proof of the renormalization of the dynamical theory obtained from the integration of Langevin equation §3.3 and the argument for the absence of power divergence in the variance of the coefficient of the perturbative expansion computed with standard NSPT §3.5.2; the generalization of the proof of convergence of KSPT to the GHMD §4.4; the implementation in `C` language of all the different NSPT algorithms discussed and the gamma method used in the analysis of the data; the generation of data and their analysis presented in §6; the developing and the implementation of the renormalization procedure Appendix: A.

Chapter 2

Lattice φ^4 theory

φ^4 theory, with φ a single-component real field, defined on a four-dimensional Euclidean lattice of extent L in all directions, is specified by the lattice action

$$S(\varphi) = a^4 \sum_{x\in\Omega} \left(\frac{1}{2}\varphi(x)\Delta\varphi(x) + \frac{1}{2}m_0^2\varphi(x)^2 + \frac{g_0}{4!}\varphi(x)^4 \right), \tag{2.0.1}$$

where φ is the *bare* field, $\Delta\varphi(x) = \big(\varphi(x+a\hat{\mu}) - 2\varphi(x) + \varphi(x-a\hat{\mu})\big)/a^2$ is the lattice Laplacian with $\hat{\mu}$ being a unit vector in the direction $\mu = 0, \ldots, 3$, and a is the lattice spacing. The sum in (2.0.1) runs over the set Ω of all lattice points $x = (x_0, x_1, x_2, x_3)$ with $x_i/a \in \mathbb{Z}_{L/a}$, while the field φ satisfies the periodicity conditions $\varphi(x + \hat{\mu}L) = \varphi(x)$, $\forall\mu$. The parameters m_0 and g_0 are the *bare* mass and coupling constant; they are related to the renormalized quantities m and g by

$$m^2 = m_0^2 - \delta m^2 = m_0^2 - \sum_{k=1}^{\infty} m_k^2\, g_0^k, \tag{2.0.2}$$

$$g = g_0 - \delta g = g_0 + \sum_{k=2}^{\infty} c_k\, g_0^k, \tag{2.0.3}$$

where the coefficients m_k^2 and c_k of the mass and coupling counterterms δm^2 and δg are determined order by order in the coupling from the renormalization conditions.

Given these definitions the expectation value of a generic observable $\mathcal{O}(\varphi)$ of the field is defined, as usual, through the Euclidean functional integral

$$\langle \mathcal{O} \rangle = \frac{1}{\mathcal{Z}} \int D\varphi \, e^{-S(\varphi)} \mathcal{O}(\varphi), \quad D\varphi \equiv \prod_{x \in \Omega} \mathrm{d}\varphi(x), \tag{2.0.4}$$

where the constant $\mathcal{Z}$ is fixed by the condition $\langle 1 \rangle = 1$. Of interest for the following discussion is the bare two-point function,

$$\chi_2(p) = a^4 \sum_{x \in \Omega} e^{-ip \cdot x} \big\langle \varphi(x) \varphi(0) \big\rangle, \tag{2.0.5}$$

where $p = (p_0, p_1, p_2, p_3)$ with $p_i = 2\pi n_i / L$ and $n_i \in \mathbb{Z}_{L/a}$ are the allowed momenta in a periodic box; the set of such momenta will be denoted in the following by $\tilde{\Omega}$

$$\tilde{\Omega} = \left\{ p = (p_0, p_1, p_2, p_3) \,|\, p_i = 2\pi n_i / L \,\text{with}\, n_i \in \mathbb{Z}_{L/a} \right\}. \tag{2.0.6}$$

In particular, we will consider

$$\chi_2 \equiv \chi_2(0) \quad \text{and} \quad \chi_2^* \equiv \chi_2(p_*), \tag{2.0.7}$$

where p_* is the minimal non-zero momentum given by $p_* = (2\pi/L, 0, 0, 0)$.[1]

2.1 Renormalization conditions and observables

In order to study the continuum limit of the theory some renormalization conditions must be chosen to define the renormalized parameters and fields; we use the finite size renormalization scheme described in [17].

Our definition of a renormalized mass m is obtained from

$$\frac{\chi_2}{\chi_2^*} = 1 + \frac{\hat{p}_*^2}{m^2}, \tag{2.1.1}$$

[1] In general we shall consider lattice units where $a = 1$ from now on. Nevertheless, the lattice spacing may be included in some formulas for clarity.

where $\hat{p}^2 = \sum_\mu \hat{p}_\mu^2$, with $\hat{p}_\mu = 2\sin(p_\mu/2)$ being the usual lattice momenta. The wavefunction renormalization $Z = Z(g_0, L/a, am)$ which defines the renormalized elementary field $\varphi_R(x) = Z^{-1/2}\varphi(x)$ is fixed by

$$Z^{-1} = \frac{\chi_2^{*\,-1} - \chi_2^{-1}}{\hat{p}_*^2} \quad \Longrightarrow \quad Z = m^2\chi_2. \tag{2.1.2}$$

Our renormalization conditions (2.1.1) and (2.1.2) correspond to choose δm and Z such that

$$\chi_2(0) = \frac{1}{m^2} \quad \text{and} \quad \chi(p_*^2) = \frac{1}{p_*^2 + m^2}. \tag{2.1.3}$$

This renormalization scheme differs from the more traditional on-shell scheme (see chapter 10 of [18]) which is defined by the condition

$$\chi(p^2) = \frac{1}{p^2 + m_{phys.}^2} + (\text{term regular at } p^2 = m_{phys.}^2.) \tag{2.1.4}$$

The latter equation contains two condition, specifying both the location of the pole at $m_{phys.}$ and its residue. The renormalized mass $m_{phys.}$ (2.1.4) is the physical one while m (2.1.3) it is not, however in this work we decide to use the non physical renormalization conditions (2.1.3) since they are easier to implement in a lattice simulation. A connection between the two schemes can be done considering the two-point function of the full theory

$$\chi_2(p) = \frac{1}{p^2 + m_{phys.}^2 + \Sigma(p^2)}, \tag{2.1.5}$$

the conditions (2.1.3) implies

$$m_{phys.}^2 + \Sigma(0) = m^2 \quad \text{and} \quad m_{phys.}^2 + \Sigma(p_*^2) = m^2\,, \tag{2.1.6}$$

while the condition (2.1.4) implies

$$\Sigma(m_{phys.}^2) = 0 \quad \text{and} \quad \frac{\partial\Sigma}{\partial p^2}(m_{phys.}^2) = 0\,. \tag{2.1.7}$$

Equations (2.1.6) relates the two different renormalization scheme.

The finite size continuum limit may then be defined by keeping the combination

$$z = mL \tag{2.1.8}$$

fixed.

More precisely, for a given choice of z the continuum limit is approached by taking the lattice size $L = L/a \to \infty$ while tuning the lattice mass $m = am \to 0$, such that z has the desired value. The possible values of z thus identify a *family* of renormalization schemes.

Given these definitions, we introduce the renormalized coupling

$$g = -\frac{\chi_4}{\chi_2^2}\, m^4, \tag{2.1.9}$$

where χ_4 is the bare connected four-point function at zero external momenta,

$$\chi_4 = \sum_{x,y,z\in\Omega} \big\langle \varphi(x)\varphi(y)\varphi(z)\varphi(0)\big\rangle - 3L^4\chi_2^2. \tag{2.1.10}$$

The coupling (2.1.9) is known to two-loop order in lattice perturbation theory [17]: On the other hand, a precise determination of (2.1.10) using the Monte Carlo methods presented in the next chapters is difficult on large lattices (required to be close to the continuum limit) due to the stochastic subtraction of the disconnected contribution.

In order to obtain precise and simple quantities with well-defined continuum limits we consider observables defined through the gradient flow [19, 20]. In the case of the φ^4 theory the gradient flow equations take the simple form [21, 22],

$$\partial_t\tilde{\varphi}(t,x) = \Delta\tilde{\varphi}(t,x) \quad \text{with} \quad \tilde{\varphi}(0,x) = \varphi_R(x), \tag{2.1.11}$$

where $t \geq 0$ is the flow time and $\Delta\varphi(x) = \big(\varphi(x+a\hat{\mu}) - 2\varphi(x) + \varphi(x-a\hat{\mu})\big)$, being the usual lattice Laplacian. In particular, we will see in §2.2, products of fields at positive flow time are automatically renormalized if the parameters of the theory

are renormalized. The dimensionless quantity

$$\mathcal{E}(t) = t^2\langle E(t,x)\rangle \quad \text{with} \quad E(t,x) = \tilde{\varphi}(t,x)^4, \tag{2.1.12}$$

for example, is finite without any additional renormalization, provided that the physical flow time t is held fixed as the continuum limit of the theory is approached (see §2.2). We define the finite size continuum limit of flow quantities like (2.1.12) by holding the ratio [23]

$$c = \sqrt{8t}/L \tag{2.1.13}$$

fixed. The continuum limit is thus taken by increasing the lattice size $L = L/a$ and the flow time in lattice units $t = t/a^2$ such that c is fixed to some chosen value; different values of c define different renormalization schemes.

2.2 Gradient flow effect in the composite operator renormalization

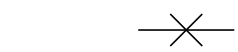
Here we illustrate how the gradient flow, introduced in the previous section, eliminates the necessity for further renormalization in composite operators besides the one of the mass (2.1.1), wavefunction (2.1.2) and coupling (2.1.9). In order to give a simpler description of the effects of the gradient flow we consider as an example the renormalization of the φ^2 operator in six dimensional φ^3 theory. This is because in φ^4 theory there is no wavefunction renormalization at one loop, so the gradient flow effects are completely manifest at two loops. Further, to simplify more the problem we will work in dimensional regularization, where respect to the lattice regularization it is easier extract the divergence.

First we illustrate how to renormalize the φ^2 operator in φ^3 theory in six dimension without the gradient flow. The analysis of chapter 6.2 of [24] can be done in Euclidean time. Consider an insertion of the φ^2 operator in the correlation function

$$\langle\varphi(x)\varphi(y)\varphi^2(z)/2\rangle \tag{2.2.1}$$

We represent the vertex of $\varphi^2(z)/2$ with a cross and the associated Feynman rule

is



$$= 1 \tag{2.2.2}$$

The connected diagrams of (2.2.1) up to g^2 order are illustrated in figure 2.1

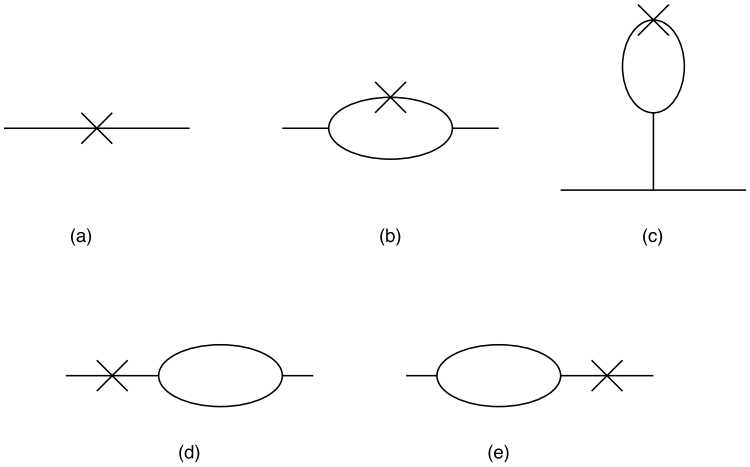


Figure 2.1 *Feynman diagram of* 2.2.1 *up to* g^2 *order.*

Working in momentum space we have

$$G = \langle \tilde{\varphi}(p_1)\tilde{\varphi}(p_2)\varphi^2(z)/2\rangle = \int d^dx\, d^dy e^{i(p_1x+p_2y)}\langle \varphi(x)\varphi(y)\varphi^2(z)/2\rangle. \tag{2.2.3}$$

The graph of order g^0, figure 2.1(a), is equal to

$$G_a = \frac{1}{(p_1^2+m^2)}\frac{1}{(p_2^2+m^2)}. \tag{2.2.4}$$

The one loop diagrams figure 2.1(d) and (e) are divergent, their divergence is removed by the inclusion of the diagrams with the wavefunction and mass counterterm figure 2.2.

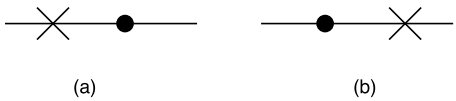


Figure 2.2 *Graph with the counterterms for figure 2.1(d) and (e). The heavy dot represent the counterterms.*

The other two graph (b) and (c) of figure 2.1 have no counterterms and are

both divergent. For figure 2.1 (b) in dimensional regularization $d = 6 - \epsilon$ we have

$$G_b = \frac{1}{(p_1^2 + m^2)} \frac{1}{(p_2^2 + m^2)} \times \frac{-g^2 \mu^{6-d}}{(2\pi)^d} \int d^d k \frac{1}{(k^2 + m^2)[(k - p_1)^2 + m^2][(k + p_2)^2 + m^2]} \tag{2.2.5}$$

and for 2.1 (c)

$$G_c = \frac{-g\mu^{3-d/2}}{(p_1^2 + m^2)(p_2^2 + m^2)[(p_1 + p_2)^2 + m^2)]} \times \frac{g\mu^{3-d/2}}{2(2\pi)^d} \int d^d k \frac{1}{(k^2 + m^2)[(k + p_1 + p_2)^2 + m^2]}. \tag{2.2.6}$$

The fact that the sum $G_b + G_c$ diverges means that the operator φ^2 is not finite. In order to have a finite operator we have to define the Zimmermann's normal operator $N(\phi^2)$ [24]. This is to be expected from power-counting.

$$\frac{1}{2} N(\varphi^2) \equiv Z_a \frac{1}{2} \varphi^2 + \mu^{d/2-3} Z_b m^2 \varphi + \mu^{d/2-3} Z_c \partial_\mu \partial^\mu \varphi. \tag{2.2.7}$$

The operators φ and $\partial_\mu \partial^\mu \varphi$ are said to mix with φ^2 under renormalization. The renormalization factor using minimal subtraction are calculated in chapter 6.2 of [24]

$$Z_a = 1 + \frac{g^2}{64\pi^3(d-6)} + O(g^3) \tag{2.2.8}$$

$$Z_b = \frac{g}{64\pi^3(d-6)} + O(g^2) \tag{2.2.9}$$

$$Z_c = \frac{g}{64\pi^3(d-6)6} + O(g^2). \tag{2.2.10}$$

The operator φ and $\partial_\mu \partial^\mu \varphi$ are finite thus φ^2, φ and $\partial_\mu \partial^\mu \varphi$ form a closed set under renormalization.

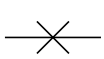
Now we will consider the case of an insertion of the operator $\varphi^2(t, x)$ at positive flow time, in the correlator

$$\langle \varphi(x) \varphi(y) \varphi^2(t, z)/2 \rangle. \tag{2.2.11}$$

The gradient flow equation (2.1.11) can be solved in momentum space

$$\tilde{\varphi}(t,p) = e^{-p^2 t}\tilde{\varphi}(0,p) \quad \text{with} \quad \tilde{\varphi}(0,p) = \int d^d x\, e^{ipx}\varphi(x). \tag{2.2.12}$$

The Feynman rule associated with an insertion of $\phi^2(t,x)/2$ will change from (2.2.2) to

 $= e^{-(p_1^2+p_2^2)t}$. (2.2.13)

We can see that the divergent part of figure 2.1 (d) and (e) is the same of the one at zero flow time, because the flow vertex is outside the loop in both case. Thus the diagrams in figure 2.1 (d) and (e) are finite after adding the diagrams with counterterms (figure 2.2). Instead for figure 2.1 (b) at positive flow time we have

$$\begin{aligned} G_b(t) &= \frac{1}{(p_1^2+m^2)}\frac{1}{(p_2^2+m^2)} \\ \times\frac{-g^2\mu^{6-d}}{(2\pi)^d}\int d^d k &\frac{e^{-(k-p_1)^2 t}e^{-(k+p_2)^2 t}}{(k^2+m^2)[(k-p_1)^2+m^2][(k+p_2)^2+m^2]} \end{aligned} \tag{2.2.14}$$

and for 2.1 (c)

$$\begin{aligned} G_c &= \frac{-g\mu^{3-d/2}}{(p_1^2+m^2)(p_2^2+m^2)[(p_1+p_2)^2+m^2)]} \\ \times\frac{g\mu^{3-d/2}}{2(2\pi)^d}\int d^d k &\frac{e^{-k^2 t}e^{-(k+p_1+p_2)^2 t}}{(k^2+m^2)[(k+p_1+p_2)^2+m^2]}. \end{aligned} \tag{2.2.15}$$

We can see that the integral in $G_b(t)$ and $G_c(t)$ are damped by exponentials of the flow time and so they are finite. Thus at positive flow time the operator $\varphi^2(t,x)$ is finite after mass, coupling and wavefunction renormalization. The same argument applies without any additional difficulties to the Euclidean Lattice formulation of the theory

2.3 Lattice perturbation theory result

In the previous section we argue that at positive flow time a composite operator does not require extra renormalization besides the condition 2.1.1 2.1.2 and 2.1.9.

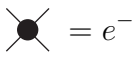
Here we will compute the first orders in the perturbative expansion of the operator (2.1.12) in lattice ϕ^4 theory. The data obtained will be used as a comparison to check the correctness of the numerical stochastic perturbation theory methods that will be discussed later. As in the continuum the gradient flow equation (2.1.11) can be solved in momentum space giving

$$\tilde{\varphi}(t,p) = e^{-\hat{p}^2 t}\tilde{\varphi}(0,p) \tag{2.3.1}$$

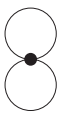
where the discrete Fourier Transform of the field is defined as

$$\tilde{\varphi}(t,p) = \frac{1}{\sqrt{V}}\sum_{x\in\Omega} e^{-ip\cdot x}\phi(t,x) \tag{2.3.2}$$

with $p \in \tilde{\Omega}$ (2.0.6). An insertion of the operator (2.1.12) can be represented in the Feynman diagram language as a vertex associated with the Feynman rule


$= e^{-(p_1^2+p_2^2+p_3^2+p_4^2)t}$ (2.3.3)

At g^0 order we have

$$\langle E(t,x)\rangle|_{g^0} = 3$$  $$= 3\left(\frac{1}{V}\sum_{p\in\tilde{\Omega}}\frac{e^{-2\hat{p}^2 t}}{\hat{p}^2+m^2}\right)^2 \tag{2.3.4}$$

To compute higher order in g_0 we need to impose the renormalization conditions. The mass counterterm for the renormalization scheme (2.1.1)(2.1.2) in terms of the bare coupling can be found in [17] [2]

$$\delta m^2 = m_1^2 g_0 + m_2^2 g_0^2 + O(g_0^3) \tag{2.3.5}$$

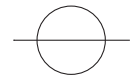
[2]Here our conversion differs from the one in [17] by a minus sign.

where

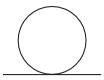
$$m_1^2 = \frac{1}{2V}\sum_{p\in\tilde{\Omega}} \frac{1}{\hat{p}^2+m^2} = \frac{1}{g}$$  , (2.3.6)

$$m_2^2 = \frac{1}{6}\left[(m^2/\hat{p}_*^2)J_2(m^2,p_*) - (1+m^2/\hat{p}_*^2)J_2(m^2,0)\right] = \frac{1}{g_0^2}$$ (2.3.7)

$J_2(m,p)$ is the value of the 1PI "sunset" diagram with external momentum p

$$J_2(m,p) = \sum_{x\in\Omega}\frac{e^{-ip\cdot x}}{V^3}\left(\sum_{q\in\tilde{\Omega}}\frac{e^{iq\cdot x}}{\hat{q}^2+m^2}\right)^3 =$$  . (2.3.8)

In this renormalization scheme the tadpole and all the graphs that contain it are cancelled by the mass renormalization at one loop

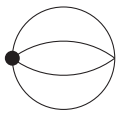
$$\frac{1}{2}$$  $$+$$ $$= 0,$$ (2.3.9)

so we will not consider these type of graphs The wavefunction and coupling renormalization do not contribute at order g_0 their first contributions are at $O(g_0^2)$

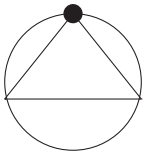
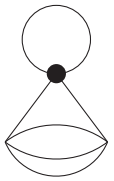
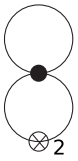
$$Z = 1 + g_0^2 Z_2 + O(g_0^3) = 1 + \frac{g_0^2}{6\hat{p}_*^2}[J_2(m^2,p_*) - J_2(m^2,0)] + O(g_0^3) \quad (2.3.10)$$

$$g = g_0 - \frac{3g_0^2}{2V}\sum_{p\in\Omega}\left(\frac{1}{\hat{p}^2+m^2}\right)^2 + O(g_0^3). \quad (2.3.11)$$

For the operator (2.1.12) at order g_0 we have

$$\langle E(t,x)\rangle|_g =$$  $$= \sum_{y\in\Omega}\left(\frac{1}{V}\sum_{p\in\tilde{\Omega}}\frac{e^{-\hat{p}^2 t+ip\cdot y}}{\hat{p}^2+m^2}\right)^4$$ (2.3.12)

At order g_0^2 the observable becomes

$$\langle E(t,x)\rangle|_{g_0^2} = \frac{3}{2}$$  $$+$$  $$+ 3$$  $$- 2\langle E\rangle|_{g^0} Z_2$$

$$= \frac{3}{2}\frac{1}{V}\sum_{p\in\tilde{\Omega}}\left[\sum_{x\in\Omega}\left(\frac{1}{V}\sum_{k\in\Omega}\frac{e^{ikx-k^2t}}{\hat{k}^2+m^2}\right)^2 e^{-ipx}\right]^2\left[\sum_{x\in\Omega}\left(\frac{1}{V}\sum_{k\in\Omega}\frac{e^{ikx}}{\hat{k}^2+m^2}\right)^2 e^{-ipx}\right]$$

$$+\left(\frac{1}{V}\sum_{k\in\Omega}\frac{e^{-2k^2t}}{\hat{k}^2+m^2}\right)\frac{1}{V}\sum_{p\in\tilde{\Omega}}\left(\frac{e^{-p^2t}}{\hat{p}^2+m^2}\right)^2\left[\sum_{x\in\Omega}\left(\frac{1}{V}\sum_{k\in\Omega}\frac{e^{ikx}}{\hat{k}^2+m^2}\right)^3 e^{-ipx}\right]$$

$$+\left(\frac{1}{V}\sum_{k\in\Omega}\frac{e^{-2k^2t}}{\hat{k}^2+m^2}\right)\frac{1}{V}\sum_{p\in\tilde{\Omega}}\left(\frac{e^{-p^2t}}{\hat{p}^2+m^2}\right)^2 m_2^2$$

$$-3\langle E\rangle|_{g^0}\frac{2}{6\hat{p}_8^2}[J_2(m^2,p)-J_2(m^2,0)]. \tag{2.3.13}$$

Using the (2.3.11) an the (2.3.13) we can obtain the coefficient of order g^2

$$\langle E(t,x)\rangle|_{g^2} = \langle E(t,x)\rangle|_{g_0^2} + \frac{3}{2V}\sum_{p\in\Omega}\left(\frac{1}{\hat{p}^2+m^2}\right)^2\langle E(t,x)\rangle|_{g^0}. \tag{2.3.14}$$

In table 2.1 we list the value of the coefficient of the expansion of $t^2\langle E\rangle$ in the bare coupling approaching the continuum limit with $z=4$.

L	$t^2\langle E\rangle\|_{g^0}$	$t^2\langle E\rangle\|_g$	$t^2\langle E\rangle\|_{g^2}$
4	4.5921×10^5	-2.3125×10^7	1.5351×10^6
6	4.7991×10^5	-2.1524×10^7	2.0758×10^6
8	3.3770×10^5	-1.0585×10^7	1.6840×10^6
10	2.5813×10^5	-6.1927×10^8	1.4138×10^6
12	2.2347×10^5	-4.7649×10^8	1.3108×10^6
14	2.0717×10^5	-4.1947×10^8	1.2818×10^6
16	1.9827×10^5	-3.9041×10^8	1.2810×10^6
18	1.9276×10^5	-3.7301×10^8	1.2914×10^6
20	1.8906×10^5	-3.6156×10^8	1.3066×10^6
22	1.8644×10^5	-3.5354×10^8	1.3239×10^6
24	1.8450×10^5	-3.4768×10^8	1.3419×10^6
26	1.8303×10^5	-3.4325×10^8	1.3600×10^6
28	1.8188×10^5	-3.3982×10^8	1.3779×10^6
30	1.8097×10^5	-3.3710×10^8	1.3954×10^6
32	1.8023×10^5	-3.3490×10^8	1.4124×10^6

Table 2.1 *Values of the first coefficient of the expansion of $t^2\langle E\rangle$ in g. The continuum limit is approched keeping $z = 4 = mL$ and $c = 0.2$. The errors on these numbers are smaller than the digit quoted.*

Chapter 3

Stochastic Quantization

Typically in quantum field theory one is interested in computing correlation functions defined as

$$\langle \varphi(x_1) \cdots \varphi(x_n) \rangle = \frac{\int \mathcal{D}\,\phi e^{-S_E} \varphi(x_1) \cdots \varphi(x_n)}{\int \mathcal{D}\,\phi e^{-S_E}} \tag{3.0.1}$$

where S_E is the euclidean action. Stochastic quantization was proposed by Parisi and Wu [14], it is a method for computation of the correlation functions(3.0.1). The main idea is to consider the Euclidean path integral as the stationary distribution of a stochastic process.
The stochastic quantization procedure is:

1. Introduce a fictitious time coordinate t_s.

2. Require that the field $\phi(t_s, x)$ is a solution of a stochastic equation, for example the Langevin equation

$$\partial_{t_s} \phi(t_s, x) = -\frac{\delta S_E[\phi]}{\delta \phi(t_s, x)} + \eta(t_s, x), \tag{3.0.2}$$

   where $\frac{\delta S_E[\phi]}{\delta \phi(t_s,x)}$ denotes the field-derivative of the euclidean action extended

in the new time dimension

$$S_E[\phi] = \int dt_s d^D x \mathcal{L}\big(\phi(t_s, x), \partial_\mu \phi(t_s, x)\big), \tag{3.0.3}$$

and field η is a Gaussian random field satisfying

$$\langle \eta(t_s, x)\rangle_\eta = 0, \qquad \langle \eta(t_s, x)\eta(t'_s, y)\rangle_\eta = 2\delta(t_s - t'_s)\delta(x - y). \tag{3.0.4}$$

3. Given an initial condition to the field ϕ at time t_{s_0} we can solve the Langevin equation (3.0.2). Calling $\phi_\eta(t_s, x)$ the solution at time t_s we can average over the gaussian field η

$$\langle \phi_\eta(t_{s_1}, x_1) \cdots \phi_\eta(t_{s_n}, x_n)\rangle_\eta = \frac{\int d\eta \exp\big\{ -\frac{1}{4}\int dt_s d^D x \eta^2(t_s, x)\big\}\phi_\eta(t_{s_1}, x_1) \cdots \phi_\eta(t_{s_n}, x_n)}{\int d\eta \exp\big\{ -\frac{1}{4}\int dt_s d^D x \eta^2(t_s, x)\big\}}. \tag{3.0.5}$$

In the limit of the $t_s \to \infty$ the correlation function in the (3.0.5) goes to the correlation function of the original theory computed with action S_E (3.0.1).

$$\lim_{t_s \to \infty} \langle \phi_\eta(t_s, x_1) \cdots \phi_\eta(t_{s_n}, x_n)\rangle_\eta = \langle \varphi(x_1) \cdots \varphi(x_n)\rangle \tag{3.0.6}$$

In order to explain why the correlation function (3.0.5) in the limit of $t_s \to \infty$ go to the (3.0.1) we first have to formalize the Langevin equation (3.0.2). Following [25] we define the Langevin equation as a stochastic differential equation of the Ito-type

$$\delta\phi(t_s, x) = \phi(t_s + dt_s, x) - \phi(t_s, x) = -\frac{\delta S_E[\phi]}{\delta\phi(t_s, x)} dt_s + dw(t_s, x) \tag{3.0.7}$$

where $dw(t_s, x)$ is defined as

$$\langle dw(t_s, x)\rangle_\eta = 0 \quad \langle dw(t_s, x) dw(t'_s, y)\rangle_\eta = \begin{cases} 0 & \text{for } t_s \neq t'_s \\ 2\delta(x - y)dt_s & \text{for } t_s = t'_s \end{cases}. \tag{3.0.8}$$

dw is related to the noise η with

$$dw(t_s, x) = \int_{t_s}^{t_s+dt_s} \eta(\tau, x) d\tau \,. \tag{3.0.9}$$

We can see that (3.0.9) satisfies (3.0.8)

$$\langle dw(t_s, x) dw(t'_s, y) \rangle_\eta = \left\langle \int_{t_s}^{t_s+dt_s} \eta(\tau, x) d\tau \int_{t_s}^{t_s+dt_s} \eta(\tau', y) d\tau' \right\rangle_\eta$$
$$= \int_{t_s}^{t_s+dt_s} \int_{t_s}^{t_s+dt_s} 2\delta(x-y)\delta(\tau-\tau') d\tau d\tau' = \int_{t_s}^{t_s+dt_s} 2\delta(x-y) d\tau = 2\delta(x-y) dt_s. \tag{3.0.10}$$

dw is introduced in order to deal with the difficulty to formulate a stochastic process with a noise that has a delta function in the stochastic time as correlator (3.0.4). We notice that the magnitude of dw is of order $\sqrt{dt_s}$. Now consider a quantity that depends on the field $G(\phi(t+dt_s, x))$ at time $t+dt_s$. We can expand it as Taylor series

$$\begin{aligned} G(\phi(t_s + dt_s, x)) &= G\left(\phi(t_s, x) + \delta\phi(t_s, x)\right) \\ &= G(\phi(t_s, x)) + \int dx' \frac{\delta G(\phi(t_s, x))}{\delta\phi(t_s, x')} \delta\phi(t_s, x') + \\ &\frac{1}{2} \int\int dx' dx'' \frac{\delta^2 G(\phi(t_s, x))}{\delta\phi(t_s, x')\delta\phi(t_s, x'')} \delta\phi(t_s, x')\delta\phi(t_s, x'') + \cdots \end{aligned} \tag{3.0.11}$$

Now we can replace $\delta\phi$ with the (3.0.7) and keep the term up to order dt_s

$$G(\phi(t_s + dt_s, x)) = \tag{3.0.12}$$
$$= G(\phi(t_s, x)) + \int dx' \frac{\delta G(\phi(t_s, x))}{\delta\phi(t_s, x')} \left[-\frac{\delta S_E[\phi]}{\delta\phi(t_s, x')} dt_s + dw(x') \right]$$
$$+ \frac{1}{2} \int\int dx' dx'' \frac{\delta^2 G(\phi(t_s, x))}{\delta\phi(t_s, x')\delta\phi(t_s, x'')} dw(t_s, x') dw(t_s, x'') + O(dt_s^{\frac{3}{2}}). \tag{3.0.13}$$

From (3.0.7) we notice that $\phi(t_s, x)$ is not correlated with $\eta(t_s)$ but only with the $\eta(t_{s_1} < t_s)$ so

$$\langle G(\phi(t_s, x))dw(t_s, y)\rangle_\eta = \langle G(\phi(t_s, x))\rangle_\eta \langle dw(t_s, y)\rangle_\eta = 0\,, \tag{3.0.14}$$

$$\left\langle \frac{\delta G(\phi(t_s, x))}{\delta\phi(t_s, x)} dw(t_s, y) \right\rangle_\eta = 0. \tag{3.0.15}$$

Taking the derivative of the average over η of G and ignoring the $O(dt_s^{\frac{3}{2}})$ term we get

$$\begin{aligned}\frac{d}{dt_s}\langle G(\phi(t_s, x))\rangle_\eta &= \lim_{dt_s\to 0} \frac{\langle G(\phi(t_s + dt_s, x))\rangle_\eta - \langle G(\phi(t_s, x))\rangle_\eta}{dt_s} \\ &= -\left\langle \int dx' \frac{\delta G(\phi(t_s, x))}{\delta\phi(t_s, x')} \frac{\delta S_E[\phi]}{\delta\phi(t_s, x')} + \frac{\delta^2 G(\phi(t_s, x))}{\delta\phi(t_s, x')^2} \right\rangle_\eta .\end{aligned} \tag{3.0.16}$$

Now we introduce the probability distribution $P(\phi(x), t_s)$ that gives the average over η

$$\langle G(\phi(t_s, x)\rangle_\eta = \int \mathcal{D}\phi P(\phi(x), t_s) G(\phi(x)). \tag{3.0.17}$$

Since we will show that $P(\phi(x), t)$ in the large stochastic time limit does not depend on the initial condition we will omit the dependence on it. The (3.0.16) implies that $P(\phi(x), t)$ has to satisfy the following equation under the appropriate boundary conditions for $P(\phi(x), t)$ (e.g., $P(\phi(x), t)$ vanishes on the boundary):

$$\frac{\partial P(\phi(x), t_s)}{\partial t_s} = \int dx' \left(\frac{\delta^2}{\delta\phi(x')^2} - \frac{\delta}{\delta\phi(x')} \frac{\delta S_E[\phi]}{\delta\phi(x')} \right) P(\phi(x), t_s). \tag{3.0.18}$$

This equation is known as the Fokker-Planck equation. We can recast it by writing

$$P(\phi(x), t_s)) = \psi(\phi(x), t_s) e^{-S_E[\phi]/2} \tag{3.0.19}$$

in to a euclidean Schrödinger-type equation

$$\frac{\partial\psi(\phi(x), t_s)}{\partial t_s} = -2\hat{H}\psi(\phi(x), t_s), \tag{3.0.20}$$

where

$$\hat{H} = \frac{1}{2}\left(-\frac{\delta}{\delta\phi(x)} + \frac{1}{2}\frac{\delta S_E[\phi]}{\delta\phi(x)}\right)\left(\frac{\delta}{\delta\phi(x)} + \frac{1}{2}\frac{\delta S_E[\phi]}{\delta\phi(x)}\right)$$
$$= -\frac{1}{2}\frac{\delta^2}{\delta\phi(x)^2} + \frac{1}{8}\left(\frac{\delta S_E[\phi]}{\delta\phi(x)}\right)^2 - \frac{1}{4}\left(\frac{\delta^2 S_E[\phi]}{\delta\phi(x)^2}\right) \quad (3.0.21)$$

is Hermitian and its eigenvalues E_n are $E_n \geq 0$ [15]. We call ψ_n the eigenfunctions of H corresponding to the eigenvalues E_n, with

$$\hat{H}\psi_n = E_n\psi_n \quad (3.0.22)$$

We can expand $\psi(\phi(x), t_s)$ in eigenfunctions of $\hat{H}$

$$\psi(\phi(x), t_s) = \sum_{n=0}^{\infty} c_n\psi_n e^{-E_n t_s}. \quad (3.0.23)$$

The eigenfunction with zero eigenvalue $E_0 = 0$ is $\phi_0 = e^{-S[\phi]/2}$. It can be verified directly that

$$\hat{H}\psi_0 = 0. \quad (3.0.24)$$

If all the other eigenfunctions have eigenvalue bigger than zero [15] we have that

$$\lim_{t_s\to\infty} \psi(\phi(x), t_s) = c_0 e^{-S[\phi]/2}. \quad (3.0.25)$$

The constant c_0 can be found imposing the normalization condition for the probability distribution

$$\int \mathcal{D}\phi\, \psi(\phi(x), t_s) = 1. \quad (3.0.26)$$

So we found that for large t_s the probability distribution of finding the field in a configuration $\phi(x)$ is

$$\lim_{t_s\to\infty} P(\phi(x), t_s) = \frac{e^{-S[\phi]}}{\int \mathcal{D}\phi e^{-S[\phi]}}. \quad (3.0.27)$$

Using the definition of the probability distribution (3.0.17) we get equation (3.0.6) which proof that for stochastic quantization

$$\lim_{t_s\to\infty}\langle\phi_\eta(t_s,x_1)\cdots\phi_\eta(t_s,x_n)\rangle_\eta=\frac{\int\mathcal{D}\phi e^{-S[\phi]}\phi(x_1)\cdots\phi(x_n)}{\int\mathcal{D}\phi e^{-S[\phi]}}=\langle\varphi(x_1)\cdots\varphi(x_n)\rangle. \tag{3.0.28}$$

3.1 Stochastic perturbation theory

Stochastic quantization not only offers an alternative method to define a quantum theory, but it also defines a new form of perturbation theory. In this section we will describe how to solve perturbatively the Langevin equation, and how to use the solution to compute the correlation functions (3.0.1) of the theory. We will also show that this procedure leads to the construction of a new object called an autocorrelation function (3.2.1): this object can be considered to be a correlation function of a $d+1$ dimensional theory and can be used to analyse the algorithms for numerical simulation based on the stochastic process §3.4.

Following [15] and [26] we shall now perturbatively solve the Langevin equation. We will discuss for simplicity a scalar theory in the continuum, defined by the Euclidean action

$$S_E[\phi]=\int d^4x\,\left[\frac{1}{2}(\partial_\mu\varphi)(\partial_\mu\varphi)+\frac{1}{2}m^2\varphi^2+\frac{g}{4!}\varphi^4\right]. \tag{3.1.1}$$

The corresponding Langevin equation is

$$\partial_{t_s}\phi(t_s,x)=(\partial^2-m^2)\phi(t_s,x)-\frac{g}{3!}\phi^3(t_s,x)+\eta(t_s,x), \tag{3.1.2}$$

with η a gaussian noise field (3.0.4).

3.1.1 Leading order solution of the Langevin equation

The leading-order ($g = 0$) equation is

$$\mathcal{D}\phi_0 = \eta \tag{3.1.3}$$

$$\mathcal{D} = \partial_{t_s} - \partial_\mu \partial_\mu + m^2. \tag{3.1.4}$$

The Green function in the five-dimensional Fourier space is

$$K(t_s, x) = \int_\omega \int_p e^{-i\omega t_s + ipx} K(\omega, p)\,, \tag{3.1.5}$$

$$K(\omega, p) = (-i\omega + p^2 + m^2)^{-1} \tag{3.1.6}$$

with the notation

$$\int_\omega = \int \frac{d\omega}{2\pi} \qquad \int_p = \int \frac{d^d p}{(2\pi)^d}. \tag{3.1.7}$$

The Green function (3.1.6) has one pole at $\omega = -i(p^2 + m^2)$. Using the residue theorem it is possible to write it in the time-momentum[1] representation

$$K(t_s, p) = \int_\omega e^{-i\omega t_s} K(\omega, p) = \theta(t_s) e^{-(p^2+m^2)t_s}. \tag{3.1.8}$$

The general solution of (3.1.3) can be written

$$\phi_0(t_s, p) = \int_{-\infty}^{\infty} dt_{s_1} K(t_s - t_{s_1}, p)\eta(t_{s_1}, p) + c\; e^{-(p^2+m^2)(t_s - t_{s_0})} \tag{3.1.9}$$

where the second term is the solution of the homogeneous equation in the time-momentum representation. The constant c has to be determined by imposing some initial condition

$$\phi(t_{s_0}, p) = \phi_{in}(p). \tag{3.1.10}$$

Thus we obtain

$$c = \phi_{in}(p) - \int_{-\infty}^{\infty} dt_{s_1} K(t_{s_0} - t_{s_1}, p)\eta(t_{s_1}, p). \tag{3.1.11}$$

[1] In this case time refers to the fifth dimension, i.e., the Langevin time

Substituting the above equation into the general solution (3.1.9) we get

$$\phi_0(t_s,p) = \int_{-\infty}^{\infty} dt_{s_1} \left[K(t_s - t_{s_1},p) - K(t_{s_0} - t_{s_1},p)e^{-(p^2+m^2)(t_s-t_{s_0})}\right] \eta(t_{s_1},p) + \phi_{in}(p)e^{-(p^2+m^2)(t_s-t_{s_0})} \quad (3.1.12)$$

so that, using the θ-function inside the green function (3.1.8),we obtain for $t_s \geq t_{s_0}$

$$\phi_0(t_s,p) = \int_{t_{s_0}}^{t_s} dt_{s_1} K(t_s - t_{s_1},p)\eta(t_{s_1},p) + \phi_{in}(p)e^{-(p^2+m^2)(t_s-t_{s_0})}. \quad (3.1.13)$$

The dependence on the initial solution falls exponentially with increasing time. Since we are interested in the behaviour of the field after thermalization we can move the initial condition at $t_{s_0} \to -\infty$ obtaining

$$\phi_0(t_s,p) = \int_{-\infty}^{t_s} dt_{s_1} K(t_s - t_{s_1},p)\eta(t_{s_1},p). \quad (3.1.14)$$

3.1.2 Iterative solution of the Langevin equation

Equation (3.1.2) may be written in the form

$$\mathcal{D}\phi = \eta - \frac{g}{3!}\phi^3, \quad (3.1.15)$$

that with the use of the green function (3.1.6) may be recast as an integral equation

$$\phi(t,x) = \phi_0(t,x) - \frac{g}{3!}\int_{-\infty}^{t} ds \int d^d y K(t-s,x-y)\phi(s,y)^3\,, \quad (3.1.16)$$

that can be solved order by order by iterations. Substituting all the right hand side of the (3.1.16) to the ϕ^3 term of the same equation we get

$$\phi(t,x) = \phi_0(t,x) - \frac{g}{3!}\int_{-\infty}^{t} ds \int d^d y K(t-s,x-y)\phi_0(s,y)^3$$
$$+\frac{g^2}{2}\int_{-\infty}^{t} ds \int d^d y K(t-s,x-y)\phi_0(s,y)\phi_0(s,y)\int_{-\infty}^{t} ds' \int d^d y' K(s-s',y-y')\phi(s',y')^3 + O(g^3). \tag{3.1.17}$$

A second iteration gives

$$\phi(t,x) = \phi_0(t,x) - \frac{g}{3!}\int_{-\infty}^{t} ds \int d^d y K(t-s,x-y)\phi_0(s,y)^3$$
$$+\frac{g^2}{2}\int_{-\infty}^{t} ds \int d^d y K(t-s,x-y)\phi_0(s,y)\phi_0(s,y)\int_{-\infty}^{t} ds' \int d^d y' K(s-s',y-y')\phi_0(s',y')^3 + O(g^3). \tag{3.1.18}$$

In the frequency-momentum representations

$$\phi(\omega,p) = \int dt_{t_s} d^d x\, e^{i\omega t - ipx}\phi(t_s,x)\,, \tag{3.1.19}$$

the lowest order solution of the Langevin equation is given by

$$\phi_0(\omega,p) = K(\omega,\phi)\phi_0(\omega,p)\,. \tag{3.1.20}$$

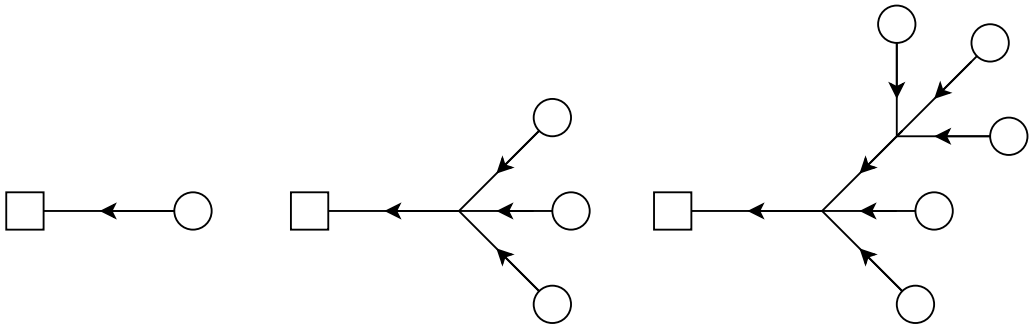


Figure 3.1 *Tree diagrams representing the solution of (3.1.16) in perturbation theory up to second order in frequency momentum representation. All diagrams have a single external line attached to the square vertex with frequency ω and momentum p going from the square to the external circles.*

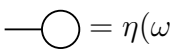
Equation (3.1.2) in frequency momentum representation becomes

$$\phi(\omega_0,p_0)=\phi_0(\omega_0,p_0)$$
$$-\frac{g}{3!}\int_{\substack{\omega_1,p_1\\ \omega_2,p_2\\ \omega_3,p_3}}\left[(2\pi)^{d+1}\delta\left(\sum_{i=0}^{3}\omega_i\right)\delta\left(\sum_{i=0}^{3}p_i\right)K(\omega,p)\phi_0(\omega_1,p_1)\phi_0(\omega_2,p_2)\phi_0(\omega_3,p_3)\right]$$
$$+\frac{g^2}{2}\int_{\substack{\omega_1,p_1\\ \omega_2,p_2\\ \omega_3,p_3}}\int_{\substack{\omega_4,p_4\\ \omega_5,p_5\\ \omega_6,p_6}}\left[(2\pi)^{d+1}\delta\left(\sum_{i=0}^{3}\omega_i\right)\delta\left(\sum_{i=0}^{3}p_i\right)K(\omega,p)\phi_0(\omega_1,p_1)\phi_0(\omega_2,p_2)\right. \tag{3.1.21}$$
$$\left.(2\pi)^{d+1}\delta\left(\sum_{i=3}^{6}\omega_i\right)\delta\left(\sum_{i=3}^{6}p_i\right)K(\omega_3,p_3)\phi_0(\omega_4,p_4)\phi_0(\omega_5,p_5)\phi_0(\omega_6,p_6)\right]+O(g^3). \tag{3.1.22}$$

Each term in the expansion in g may be represented by a tree diagram with four-point vertices and one-point vertices connected by directed lines; for example ϕ_0 is represented by the left diagram of (fig. 3.1), the term proportional to g is represented by the middle diagram of (fig. 3.1), the g^2 term by the right diagram of (fig. 3.1). More precise the directed line represents the Green function

$$\longleftarrow\ =K(\omega,p)=\frac{1}{-i\omega+p^2+m^2}\,, \tag{3.1.23}$$

and the one point vertex represents a gaussian noise insertion



$$= \eta(\omega, p). \tag{3.1.24}$$

For each four-point vertex

$$= -\frac{g}{3!} \tag{3.1.25}$$

there is conservation of frequency and momentum

$$(2\pi)^{d+1}\delta(\omega_1 + \omega_2 + \omega_3 + \omega_4)\delta(p_1 + p_2 + p_2 + p_3) \tag{3.1.26}$$

with incoming frequencies and momenta.

3.2 Autocorrelation function

Once we have calculated the perturbative solution of the Langevin equation we can use it to compute the average $\langle\rangle_\eta$ over the noise η of the product $\phi(t_{s_1}, x_1)\cdots\phi(t_{s_n}, x_n)$, this object is called the n-point autocorrelation function[2]. Thus we can compute the autocorrelation function

$$A(\omega_1, p_1; \cdots; \omega_n, p_n) = \langle\phi(\omega_1, p_1)\cdots\phi(\omega_n, p_n)\rangle_\eta \tag{3.2.1}$$

in powers of the coupling g. The autocorrelation functions are correlation function of a $d+1$ dimensional theory. The correlation functions of the original theory in momentum space are obtained computing the equal time autocorrelation function

$$\langle\varphi(p_1)\cdots\varphi(p_n)\rangle = \langle\phi(t_s = 0, p_1)\cdots\varphi(t_s = 0, p_n)\rangle_\eta = \int_{\omega_1\cdots\omega_n} A(\omega_1, p_1; \cdots; \omega_n, p_n). \tag{3.2.2}$$

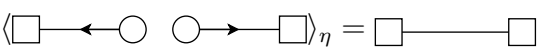
Note that there is no need to take the limit $t_s \to \infty$ since the initial condition is at $-\infty$. Contracting the noise η in the (3.2.1) using (3.0.4) one obtain a sum of

[2]The definition of autocorrelation function is the time average of $\phi(t_{s_1}, x_1)\cdots\phi(t_{s_n}, x_n)$ at fixed time separations $t_i - t_j$. The translation symmetry in time allows us to replace the time average with the average respect to the noise.

Feynman diagrams for the autocorrelation functions.

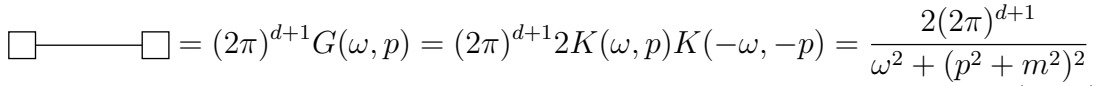
3.2.1 Feynman rules for autocorrelation functions

A noise field is represented by a one-point vertex, the contraction over two noise field is represented by an undirected line



(3.2.3)

with the Feynman rule



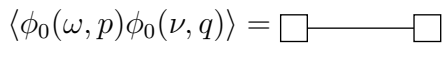
(3.2.4)

The Feynman diagrams for autocorrelations thus involve directed lines (3.1.23), undirected lines (3.2.4), and vertices (3.1.25). They are the same of an ordinary theory except for the following extra rules:

- Both directed and undirected lines can be attached to vertices;
- Each vertex has exactly one outward directed line;
- there are no diagrams with loops of directed lines;
- external lines may be only undirected lines or outward directed lines.

Each loop is associated with an integral over frequency and momentum

As an example, we compute the lowest-order two point autocorrelation function



$$= 2(2\pi)^{d+1}\delta(\omega+\nu)\delta(p+q)K(\omega,p)K(\nu,q) = \frac{2(2\pi)^{d+1}\delta(\omega+\nu)\delta(p+q)}{\omega^2+(p^2+m^2)^2}. \quad (3.2.5)$$

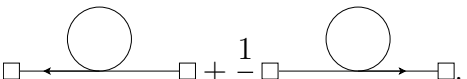
Taking the equal time autocorrelation function we obtain the correlation function of the four-dimensional theory

$$\langle\varphi(p)\varphi(q)\rangle\big|_{g^0} = \int_{\omega,\nu}\langle\phi_0(\omega,p)\phi_0(\nu,q)\rangle = \\ = \int_{\omega}\frac{2(2\pi)^d\delta(p+q)}{\omega^2+(p^2+m^2)^2} = \frac{(2\pi)^d\delta(p+q)}{(p^2+m^2)} \quad (3.2.6)$$

3.2.2 One loop two-point function

Here we will compute the one loop correction of (3.2.5) and we will show that integrating over the external frequency we recover the one loop correction to the two-point function of the φ^4 theory. At order g the two-point function in the $d+1$ dimensional theory is given by

$$\langle\phi(\omega,p)\phi(\nu,q)\rangle\big|_g = \langle\phi_0(\omega,p)\phi_1(\nu,q)\rangle + \langle\phi_1(\omega,p)\phi_0(\nu,q)\rangle$$

$= \frac{1}{2}$  $+ \frac{1}{2}$. (3.2.7)

Using the feynman rules given in the previous section we can evaluate the diagrams

$$\langle\phi(\omega,p)\phi(\nu,q)\rangle\big|_g = \\ = \frac{1}{2}(2\pi)^{d+1}\delta(\omega+\nu)\delta(p+q)\big[K(\omega,p)\Sigma(m)G(\omega,p) + G(\omega,p)\Sigma(m)K(-\omega,p)\big], \quad (3.2.8)$$

where

$$\Sigma(m) = \int_{\omega',p'}\frac{2}{\omega^2+(p^2+m^2)^2}. \quad (3.2.9)$$

The integral on the frequencies ω' in Σ can be performed obtaining

$$\Sigma(m) = \int_{p'}\frac{1}{(p^2+m^2)}. \quad (3.2.10)$$

which is the same expression of the tadopole 1PI diagram of the φ^4 theory. Using the identity

$$K(\omega, p) + K(-\omega, p) = (p^2 + m^2)G(\omega, p) \tag{3.2.11}$$

we can rewrite the (3.2.8) as

$$\langle \phi(\omega, p)\phi(\nu, q)\rangle\big|_g = \frac{1}{2}(2\pi)^{d+1}\delta(\omega + \nu)\delta(p + q)(p^2 + m^2)G(\omega, p)^2\Sigma(m). \tag{3.2.12}$$

Integrating over the external frequencies ω and ν using

$$\int_\omega G(\omega, p)^2 = \frac{1}{(2\pi)}\frac{1}{(p^2 + m^2)^2} \tag{3.2.13}$$

we get that the (3.2.8) become

$$\langle \phi(\omega, p)\phi(\nu, q)\rangle\big|_g = \frac{1}{2}(2\pi)^d\delta(p + q)\frac{1}{(p^2 + m^2)^2}\Sigma(m), \tag{3.2.14}$$

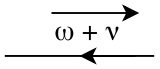
which is the one loop correction of the two point function in φ^4 theory.

3.3 Renormalization of Langevin equation

The theory constructed from the action (3.1.1) requires renormalization. So does the dynamical theory constructed from the Langevin equation. In [27] It was proved that the renormalization of the static theory together with the time-scale renormalization will render the dynamical theory finite. The approach used in [27] was to construct a functional representation of the stochastic time-dependent ϕ field correlation functions in terms of an effective local action S, which allowed the author to calculate the autocorrelations functions (3.2.1) by conventional path integral methods. Here we will show that is possible to prove the renormalization of the dynamical theory obtained from the Langevin equation using the diagrammatic expansion introduced in the previous sections. We will formulate power-counting rules for the autocorrelation function of the dynamical theory and discuss the changes that have to be done in order to apply the proof of the BPH theorem presented in [28]

3.3.1 Definitions

A Feynman integral $I(\mathcal{G})$ may be associated with any graph $\mathcal{G}$ by means of the Feynman rules for the theory, in the case of the autocorrelation function of the dynamical Langevin theory the Feynman rules are described in §3.2.1. $I(\mathcal{G})$ is a function of the external frequency ω, the momentum p, the mass m (it is assumed that $m > 0$ so as to avoid infrared divergences), a dimensionless coupling g, and a cutoff $\Lambda \equiv 1/a$ that is introduced to make the theory well-defined.
Since the propagator has the form (3.1.23) we may consider it as a function of the complex number $z = -i\omega + p^2$, in general the frequency and momentum associated with line ℓ ω_ℓ and p_ℓ are a sum of loop and external frequencies and momenta. For example, in a line of the type



$$= \frac{1}{-i(\omega+\nu)+(p+q)^2+m^2} = \frac{1}{z+m^2}\,, \qquad z = -i(\omega+\nu)+(p+q)^2 \tag{3.3.1}$$

We define $I_\lambda(\mathcal{G})$ to be the Feynman integral corresponding to $\mathcal{G}$ where all the lines have $|z|$ (The norm used here is $|z|^2 = z^* z = \omega^2 + p^4$ for $z = -i\omega + p^2$) bigger than λ, i.e., $|z_\ell| > \lambda \quad (\forall \ell \in \mathcal{G})$. $i_\lambda(\mathcal{G})$ is defined to be the integrand of the graphs $\mathcal{G}$. We define

$$\iint_{|z|>\lambda} \equiv \int_\omega \int_p \theta(|z| - \lambda) \tag{3.3.2}$$

3.3.2 R operation

We now adapt the henge decomposition defined in [28, 29] in the case of the dynamical Langevin $d+1$ theory. Any graph may be decomposed into a set of disjoint 1PI components and a set of edges which do not belong to any 1PI subgraph. Selecting any line from a graph defines a henge, which is just the set of 1PI components of the graph with the specified line removed. For examples of henges see § 2.2 of [28]. At every point in the space of loop frequencies and loop

momenta ν, k some line ℓ has to carry the smallest $|z| = |-i\omega_\ell + p_\ell^2|$:

$$I_\lambda(\mathcal{G}) = \sum_{\ell\in\mathcal{G}} \iint_{|z|>\lambda} i_{|z|}(\mathcal{G}/\mathcal{H}(\mathcal{G},\ell)) \prod_{\Theta\in\mathcal{H}(\mathcal{G},\ell)} I_{|z|}(\Theta) \tag{3.3.3}$$

where $\mathcal{H}(\mathcal{G},\ell)$ is the henge of $\mathcal{G}$ selecting the line ℓ and $\mathcal{G}/\mathcal{H}(\mathcal{G},\ell)$ is a single loop graph with the $1PI$ subgraphs in which the $\mathcal{H}(\mathcal{G},\ell)$ act as effective vertices [28, 29].

The $\bar{R}$-operation, which subtracts all subdivergences, may be defined as

$$\bar{R}I_\lambda(\mathcal{G}) \equiv \sum_{\ell\in\mathcal{G}} \iint_{|z|>\lambda} i_{|z|}(\mathcal{G}/\mathcal{H}(\mathcal{G},\ell)) \prod_{\Theta\in\mathcal{H}(\mathcal{G},\ell)} RI_{|z|}(\Theta), \qquad RI_\lambda(\mathcal{G}) \equiv \bar{R}I_\lambda(\mathcal{G}) - K\bar{R}I_0(\mathcal{G}) \tag{3.3.4}$$

where R is the operation that removes all divergences.

3.3.3 The subtraction operation -K

The subtraction operator $-K$ removes the divergent part of $I(\mathcal{G})$. It can be choose to be a Taylor series subtraction operator with respect to the external frequency ω and momentum p $T_\omega^\alpha T_p^\beta I(\mathcal{G})$ where $\alpha + 2\beta = \deg \mathcal{G}$. The $(deg\,\mathcal{G})$ is the overall (power counting) degree of divergence of $\mathcal{G}$, whose precise definition will be given in §3.3.7. The subtraction $-K$ is defined to replace the divergent part with a finite polynomial in the external frequency and momenta. The renormalization conditions specify unambiguously the finite part of a subtracted graph, so they fix the value of

$$I(p_0,\omega_0), \partial_p I(p_0,\omega_0), \partial_\omega I(p_0,\omega_0), \cdots, \partial_\omega^\alpha \partial_p^\beta I(p_0,\omega_0)$$

at the subtraction point p_0, ω_0, with $\alpha + \beta/2 = deg\,\mathcal{G}$.

3.3.4 Example of the R subtraction: one loop

In order to clarify how the R operation removes the divergences of a graph we consider the following divergent Feynman integral with dimensional regularization in $d = 4 - \epsilon$ dimensions

$$I(s) = \int_k \frac{1}{k^2 + m^2} \frac{1}{(k+s)^2 + m^2} = \frac{1}{8\pi^2} \frac{1}{\epsilon} + finite\,, \tag{3.3.5}$$

where the latter integral is the one loop 1PI four-point function in the s channel (see chapter 10.2 of [18]). The R operation in this case subtract to the integrand its Taylor series in s truncated at order zero.

$$RI(s) = I(s) - KI(s) = I(s) - I(0). \tag{3.3.6}$$

With the standard technique to extract the divergent part of the Feynman diagrams (Appendix A.4 of [18]) we have

$$I(0) = \int_k \frac{1}{k^2 + m^2} \frac{1}{k^2 + m^2} = \frac{1}{8\pi^2} \frac{1}{\epsilon} + finite\,, \tag{3.3.7}$$

from which follows that $RI(s)$ is finite.

3.3.5 Tree level bounds

In order to prove the renormalization of the dynamical Langevin theory in $d+1$ dimension we have to show that it satisfies the tree level bounds:

1. Each vertex and propagator Γ in the theory has to satisfy the inequality $|I_\lambda(\Gamma)| \leq c\chi(\lambda)^{deg\,\Gamma}$ with c a constant
2. The bound function $\chi(|z|)^{deg\,\Gamma}$ must have the properties

$$\iint_{|z|\geq\lambda} \chi(|z|)^\nu \leq c\chi(\lambda)^{\nu+\frac{d}{2}+1} \quad \text{for} \quad \nu < -(\frac{d}{2}+1) \tag{3.3.8}$$

$$\iint_{|z|\leq\lambda} \chi(|z|)^\nu \leq c\chi(\lambda)^{\nu+\frac{d}{2}+1} \quad \text{for} \quad \nu \geq -(\frac{d}{2}+1) \tag{3.3.9}$$

3. Differentiation with respect the external momenta must lower the degree of divergence

$$|\partial^n I_\lambda(\Gamma)| \leq c\chi(\lambda)^{deg\,\Gamma-\alpha n} \quad \text{with} \quad \alpha \geq 1 \tag{3.3.10}$$

3.3.6 Proof of the tree level bounds

In the case of the dynamical $d+1$ theory derived by the Langevin equation (3.1.2) the vertices are constant thus their degree of divergence is zero. The bounding function for the propagator (3.1.23) can be chosen to be $\chi(z) = \max(|z|, m^2)$

$$\left|\frac{1}{-i\omega + p^2 + m^2}\right|_{|z|>\lambda} = \left|\frac{1}{\sqrt{\omega^2 + (p^2+m^2)^2}}\right|_{|z|>\lambda} \leq c\max(\lambda, m^2)^{-1} \tag{3.3.11}$$

from which we read that the degree of divergence of the propagator is -1.

The integral over the frequency and the momenta of a d-dimensionally spherically symmetric integrand can be rewritten as

$$\begin{aligned}\iint_{|z|>\lambda} &= \int_{|z|>\lambda} d\omega dp^d = S_{d-1}\int_{|z|>\lambda} d\omega dp\, p^{d-1} = \frac{S_{d-1}}{2}\int_{|z|>\lambda} d\omega dq\, q^{d/2-1} \\ &= \frac{S_{d-1}}{2}\int_\lambda^\infty d|z|\int_0^\pi d\theta |z|(|z|\sin\theta)^{d/2-1} = c\int_\lambda^\infty d|z||z|^{d/2}\end{aligned} \tag{3.3.12}$$

where in the first step we move to spherical coordinates for the momenta, in the second step we make the change of variable $q = p^2$, and in the third step $|z|^2 = \omega^2 + q^2$, S_{d-1} is the surface area of a d-dimensional sphere, and $q = |z|\sin\theta$. If $\nu < -(d/2+1)$ the integral of the bounding function in the ultraviolet region in the case $\lambda > m^2$ is

$$\iint_{|z|>\lambda} \chi^\nu \leq c'\int_\lambda^\infty d|z||z|^{d/2}|z|^\nu \leq c|z|^{\nu+d/2+1} \quad \text{for} \quad \nu + d/2 + 1 < 0. \tag{3.3.13}$$

In the other case $\lambda < m^2$

$$\iint_{|z|>\lambda} \chi^\nu \leq c'' \left(\int_\lambda^{m^2} d|z||z|^{d/2} m^{2\nu} + \int_{m^2}^\infty d|z||z|^{d/2}|z|^\nu \right)$$
$$\leq cm^{2\nu}(m^{d+2} - \lambda^{d/2+1}) + c'm^{2\nu+d+2} \leq c(m^2)^{\nu+d/2+1} \quad \text{for} \quad \nu + d/2 + 1 < 0. \tag{3.3.14}$$

The (3.3.13) and (3.3.14) implies that the bounding function $\chi(z) = \max(|z|, m^2)$ satisfied the condition (3.3.8).

If $\nu \geq -(d/2+1)$ the integral of the bounding function in the infrared region in the case $\lambda < m$ is

$$\iint_{|z|>\lambda} \chi^\nu \leq \int_0^\lambda d|z||z|^{d/2} m^{2\nu} \leq \int_0^{m^2} d|z||z|^{d/2} m^{2\nu} \leq c(m^2)^{\nu+d/2+1} \tag{3.3.15}$$

In the other case $m < \lambda$

$$\iint_{|z|>\lambda} \chi^\nu \leq \int_0^{m^2} d|z||z|^{d/2} m^{2\nu} + \int_{m^2}^\lambda d|z||z|^{d/2}|z|^\nu \leq c\lambda^{\nu+d/2+1}. \tag{3.3.16}$$

The (3.3.15) and (3.3.16) implies that the bounding function $\chi(z) = \max(|z|, m)$ satisfied the condition (3.3.9).

The next condition to be checked is the (3.3.10). The derivative with respect to an external frequency ω gives

$$\left| \partial_\omega^n \frac{1}{-i(\omega+\nu) + (p+k)^2 + m^2} \right|_{|z|>\lambda} = \left| \frac{1}{(-i(\omega+\nu) + (p+k)^2 + m^2)^{n+1}} \right|_{|z|>\lambda}$$
$$\leq c \max(\lambda, m^2)^{-1-n} \tag{3.3.17}$$

where ν and k are loop frequency and momenta respectively, p an external momentum, ω an external frequency and $|z| = |-i(\omega+\nu) + (p+k)^2|$. The latter equation tells us that a derivative with respect to ω lowers the degree of divergence by at least one.

One derivative with respect to an external momentum p gives

$$\left|\partial_{p_i}\frac{1}{-i(\omega+\nu)+(p+k)^2+m^2}\right|_{|z|>\lambda} = \left|\frac{2(p_i+k_i)}{(-i(\omega+\nu)+(p+k)^2+m^2)^2}\right|_{|z|>\lambda}$$
$$\leq \frac{2\sqrt{(p+k)^2}}{[(\omega+\nu)^2+((p+k)^2+m^2)^2]}\Bigg|_{|z|>\lambda} \leq \frac{2\left[((p+k)^2+m^2)^2\right]^{1/4}}{[(\omega+\nu)^2+((p+k)^2+m^2)^2]}\Bigg|_{|z|>\lambda}$$
$$\leq \frac{2\left[(\omega+\nu)^2+((p+k)^2+m^2)^2\right]^{1/4}}{[(\omega+\nu)^2+((p+k)^2+m^2)^2]}\Bigg|_{|z|>\lambda} \leq \frac{2}{|-i(\omega+\nu)+((p+k)^2+m^2)|^{1+1/2}}\Bigg|_{|z|>\lambda}$$
$$\leq c\max(\lambda,m^2)^{-1-1/2}. \tag{3.3.18}$$

With similar steps it is possible to show that further derivative with respect to p decrease the degree of divergence by $1/2$ each.

Since an undirected line (3.2.4) is the product of two $K(\omega,p)$ we can bound it as $G(\omega,p)\big|_{|z|>\lambda} = \max(\lambda,p)^{-2}$, which means that it has a degree of divergence equal to -2, a derivative with respect to ω will lower the degree of divergence by one and a derivative respect p will lower the degree of divergence by $1/2$

3.3.7 Power counting in the dynamical Langevin $d+1$ theory

In the case of the dynamical Langevin $d+1$ theory the power counting can be chosen as follows:

- A loop (3.3.12) has degree of divergence $d/2+1$;
- A direct line (3.1.23) has degree of divergence $deg\, I_K = -1$;
- An undirected line (3.2.4) has degree of divergence $deg\, I_G = -2$;
- n derivatives of a line with respect to ω (3.3.17) decrease the degree of divergence to $deg\, \partial^n_\omega I_K = -1-n$ or $deg\, \partial^n_\omega I_G = -2-n$;
- n derivatives of a line with respect to p (3.3.18) decrease the degree of divergence to $deg\, \partial^n_p I = -1-n/2$ or $deg\, \partial^n_p I_G = -2-n/2$;

- The four point vertex has degree of divergence $deg\, V = 0$.

All the Feynman diagram can be decomposed in to one particle irreducible (1PI) parts and the lines connecting them. The total degree of divergence of a Feynman integral $I(\mathcal{G})$ associated to the 1PI graph $\mathcal{G}$ with L loops, I_K internal directed line, I_G internal undirected line and V vertices is

$$deg\, \mathcal{G} \equiv L(\frac{d}{2} + 1) + I_K(deg\, I_K) + I_G(deg\, I_G) + V(deg\, V). \tag{3.3.19}$$

With the given definitions of the henge decomposition (3.3.3), degree of divergence of a graph (3.3.19) and tree level bounds (3.3.5) the proof of the BPH theorem presented in [28] applies to the dynamical Langevin $d + 1$ theory. The BPH theorem proved in [28] states that any subtracted Feynman integral of the graph $\mathcal{G}$ is convergent, i.e., they have a finite limit as the cutoff $\lambda \to \infty$. Moreover the subtractions are local polynomial of $deg\, \mathcal{G}$ in ω and $2deg\, \mathcal{G}$ in p.

For a connected graph, we require exactly $V - 1$ lines to connect V vertices into a tree; every extra line produces a loop: hence

$$L = I - V + 1 = I_K + I_G - V + 1. \tag{3.3.20}$$

Every line has to end on an appropriate vertex so

$$4V = E_K + E_G + 2I_K + 2I_G \tag{3.3.21}$$

where E_K is the number of external directed lines and E_G the number of external undirected lines. Since the the vertices must have exactly one outward directed line (see 3.2.1)

$$E_K + I_K = V. \tag{3.3.22}$$

The sum of the undirected and directed external lines has to give the total number of external lines

$$E = E_K + E_G. \tag{3.3.23}$$

Setting $d = 4$ and using (3.3.20),(3.3.21),(3.3.22) and (3.3.23) to express $deg\, \mathcal{G}$

(3.3.19) in terms of E_G and E we get

$$deg\,\mathcal{G} = 3 - \frac{3}{2}E + E_G. \tag{3.3.24}$$

For E odd the Feynman integral of the graphs is automatically zero by the contraction over the noise.

Since in the graphs of an autocorrelation function (§3.2.1) for each vertex there is exactly one outward directed line and there are not loop of directed lines every graph has at least one directed external line so $E_G < E$. The maximum degree of divergence is reached when $E_G = E - 1$ i.e.

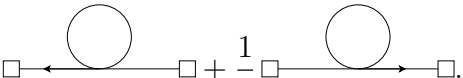
$$\max(deg\,\mathcal{G}) = 2 - \frac{1}{2}E. \tag{3.3.25}$$

The last equation tells us that only a finite number of autocorrelation functions are divergent, the two-point function with $deg\,G_2 = 1$ and the four-point function $deg\,G_4 = 0$. Since $deg\,G_2 = 1$ the divergent part of the two point function is of the form

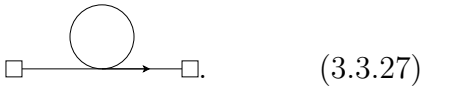
$$c_0(g) + c_1(g)\omega + c_2(g)p^2. \tag{3.3.26}$$

Lorentz invariance forbids terms linear in p. The term $c_0(g)$ correspond to the mass renormalization. The divergent part of the four-point function will be of the form $c_3(g)$. All the coefficients c_0 and c_3 starts at order g. In ϕ^4 theory the only divergent diagrams at one loop of the two-point function are the tadpoles

$$\langle\phi(\omega,p)\phi(\omega,q)\rangle_\eta\big|_g = \frac{1}{2}$$



$$+\frac{1}{2}$$



. (3.3.27)

which do not carry any external frequencies or momenta inside the loop so their divergent parts will only contribute to c_0; hence contributions to c_1, c_2 can only appear at order g^2 or higher.

It remains to be proved that all these subtractions made by the R operation sum up to counterterms in the Langevin equation. However we believe that it is so and all the divergences can be eliminated adding the following countertems on

the right side of the Langevin equation (3.1.2)

$$-(\delta m^2)m^2\phi - (\delta t_s)\partial_{t_s}\phi + (\delta Z)\partial_\mu\partial^\mu\phi - g\frac{(\delta g)}{3!}\phi^3\,, \tag{3.3.28}$$

So the renormalized Langevin equation is

$$Z_{t_s}\partial_{t_s}\phi_b(t_s,x) = (\partial^2 - Z_m m^2)\phi_b(t_s,x) - \frac{Z_g g}{3!}\phi_b^3(t_s,x) + \eta(t_s,x), \tag{3.3.29}$$

with

$$\sqrt{Z_\phi}\phi(t_s,x) = \phi_b(t_s,x) \tag{3.3.30}$$

3.4 Numerical integration

Stochastic quantization has been found to be useful for the numerical solution of quantum field theories [30] on the lattice. One starts from the Euclidean action on a d-dimension hypercubic lattice with spacing a

$$S = a^d \sum_x \big(-\frac{1}{2}\phi(x)\Delta\phi(x) + V(x)\big), \tag{3.4.1}$$

with $\Delta\phi(x) = \sum_\mu \big(\phi(x+\hat{\mu}) - 2\phi(x) + \phi(x-\hat{\mu})\big)$, with $\hat{\mu}$ being a unit vector in the direction $\mu = 0,\ldots,3$. In the case of φ^4 theory the potential is

$$V(x) = \frac{1}{2}m_0^2\phi(x)^2 - \frac{g_0}{4!}\phi^4(x) \tag{3.4.2}$$

In a numerical simulation the stochastic time t_s of the Langevin simulation corresponds to the computer time, so it is discretized in steps ϵ and we need to use a discrete form of the Langevin equations. With the Euler scheme we have

$$\partial_{t_s}\phi(x)(t_s) \to \frac{1}{\epsilon}\big(\phi(x)^{t_s+\epsilon} - \phi^{t_s}(x)\big) \tag{3.4.3}$$

and so the Langevin equation (3.0.2) become

$$\phi(x)^{t_s+\epsilon} = \phi^{t_s}(x) - \sqrt{\epsilon}\eta^{t_s}(x) - \epsilon\frac{\delta S}{\delta\phi^{t_s}(x)}, \tag{3.4.4}$$

where

$$\langle\eta^{t_s}(x)\eta^{t'_s}(y)\rangle_\eta = 2\delta_{t_s,t'_s}a^{-d}\delta_{x,y}. \tag{3.4.5}$$

It can be shown that for discrete stochastic time the equilibrium probability distribution in the limit of large t_s differs from the (3.0.27) by $O(\epsilon)$. Following [31] we will now compute the $O(\epsilon)$ term for the Euler scheme. Setting the lattice spacing $a = 1$ we call $P^{t_s}[\phi]$ the probability distribution at time t_s. After one step of (3.4.4) it goes into

$$P^{t_s+\epsilon}[\phi] = \int \mathcal{D}\phi'_x\langle\delta(\phi(x) - \phi(x)' + f_x)\rangle_\eta P^{t_s}[\phi] \tag{3.4.6}$$

where

$$f_x = \sqrt{\epsilon}\eta^{t_s}(x) + \epsilon\frac{\delta S}{\delta\phi^{t_s}(x)}. \tag{3.4.7}$$

For small ϵ we can expand the δ function in powers of f_x and integrate over ϕ' obtaining the asymptotic Kramers-Moyal expansion

$$P^{t_s+\epsilon}[\phi] = P^{t_s}[\phi] + \sum_{n=1}^{\infty}\frac{1}{n!}\frac{\delta}{\delta\phi^{t_s}(x_1)}\cdots\frac{\delta}{\delta\phi^{t_s}(x_n)}\left(\langle f_{x_1}\cdots f_{x_n}\rangle_\eta P^{t_s}[\phi]\right) \tag{3.4.8}$$

where an integral on the spatial index x_i is implied. Up to the second order we have

$$\begin{aligned}
\langle f_{x_1}\rangle_\eta &= \epsilon\frac{\delta S}{\delta\phi^{t_s}(x_1)} + O(\epsilon^3),\\
\langle f_{x_1}f_{x_2}\rangle_\eta &= 2\epsilon\delta_{x_1x_2} + \epsilon^2\frac{\delta S}{\delta\phi^{t_s}(x_1)}\frac{\delta S}{\delta\phi^{t_s}(x_2)} + O(\epsilon^3),\\
\langle f_{x_1}f_{x_2}f_{x_3}\rangle_\eta &= 2\epsilon^2\left(\delta_{x_1,x_2}\frac{\delta S}{\delta\phi^{t_s}(x_3)} + \delta_{x_2,x_3}\frac{\delta S}{\delta\phi^{t_s}(x_1)} + \delta_{x_3,x_1}\frac{\delta S}{\delta\phi^{t_s}(x_2)}\right) + O(\epsilon^3),\\
\langle f_{x_1}f_{x_2}f_{x_3}f_{x_4}\rangle_\eta &= 4\epsilon^2\left(\delta_{x_1,x_2}\delta_{x_3,x_4} + \delta_{x_1,x_3}\delta_{x_2,x_4} + \delta_{x_1,x_4}\delta_{x_3,x_2}\right) + O(\epsilon^3).
\end{aligned} \tag{3.4.9}$$

Expanding the (3.4.8) up to the first order in ϵ we get a discrete version of the

Fokker-Planck equation for the lattice action

$$\frac{1}{\epsilon}(P^{t_s+\epsilon}[\phi] - P^{t_s}[\phi]) = \frac{\delta}{\delta\phi^{t_s}(x_1)}\left(\frac{\delta S}{\delta\phi^{t_s}(x_1)} + \frac{\delta}{\delta\phi^{t_s}(x_1)}\right) P^{t_s}[\phi], \tag{3.4.10}$$

which converges to the desired equilibrium distribution $P \propto e^{-S}$. At the second order in ϵ the (3.4.8) become

$$\begin{aligned}\frac{1}{\epsilon}(P^{t_s+\epsilon}[\phi] - P^{t_s}[\phi]) = {} & \frac{\delta}{\delta\phi^{t_s}(x_1)}\left(\frac{\delta S}{\delta\phi^{t_s}(x_1)} + \frac{\delta}{\delta\phi^{t_s}(x_1)}\right) P^{t_s}[\phi] \\ & +\epsilon\left\{\frac{1}{2}\frac{\delta}{\delta\phi^{t_s}(x_1)}\frac{\delta}{\delta\phi^{t_s}(x_2)}\left(\frac{\delta S}{\delta\phi^{t_s}(x_1)}\frac{\delta S}{\delta\phi^{t_s}(x_2)}P^{t_s}[\phi]\right)\right. \\ & +\frac{\delta}{\delta\phi^{t_s}(x_1)}\frac{\delta^2}{\delta\phi^{t_s}(x_2)\delta\phi^{t_s}(x_2)}\left(\frac{\delta S}{\delta\phi^{t_s}(x_1)}P^{t_s}[\phi]\right) \\ & \left.+\frac{\delta^2}{\delta\phi^{t_s}(x_2)\delta\phi^{t_s}(x_2)}\frac{\delta^2}{\delta\phi^{t_s}(x_1)\delta\phi^{t_s}(x_1)}P^{t_s}[\phi]\right\}.\end{aligned} \tag{3.4.11}$$

Assuming that the equilibrium distribution exists and is unique, we can compute it by setting the left hand side of the above equation to zero. We can use the result of the $O(\epsilon)$ (3.4.10) at the equilibrium

$$\frac{\delta P^{t_s}}{\delta\phi^{t_s}(x_1)} = -\left(\frac{\delta S}{\delta\phi^{t_s}(x_1)}\right) P^{t_s}[\phi] + O(\epsilon), \tag{3.4.12}$$

to simplify the expression. Inserting it in the second line of the (3.4.11), we obtain

$$0 = \frac{\delta}{\delta\phi^{t_s}(x_1)}\left[\left(S + \delta S + \frac{\delta}{\delta\phi^{t_s}(x_1)}\right) P^{t_s}[\phi]\right], \tag{3.4.13}$$

where

$$\delta S = \frac{\epsilon}{4}\sum_{x_2}\left[2\frac{\delta^2 S}{\delta\phi^{t_s}(x_2)\delta\phi^{t_s}(x_2)} - \left(\frac{\delta S}{\delta\phi^{t_s}(x_2)}\right)^2\right] + O(\epsilon^2). \tag{3.4.14}$$

The above equation is a discrete version of the Fokker-Planck equation (3.0.18), the solution [15] is the equilibrium distribution $P \propto e^{-(S+\delta S)}$ that is the desired one multiplied by a factor $e^{-\delta S}$. The explicit form of the shift δS depends on the integrator.

It is possible to formulate higher order integrator schemes such that $\delta S = O(\epsilon^n)$,

In section 6 of [31] a second order integrator is given for gauges theories, in a scalar theory it will take the form of

$$\phi(x)^{t_s+\epsilon} = \phi^{t_s}(x) - \sqrt{\epsilon}\eta^{t_s}(x) - \frac{\epsilon}{a+b}\left(a\frac{\delta S}{\delta\phi}\bigg|_{\phi^{t_s}(x)} + b\frac{\delta S}{\delta\phi}\bigg|_{\phi^{a,t_s}(x)}\right) \tag{3.4.15}$$

where

$$\phi^{a,t_s}(x) = \phi^{t_s}(x) - \frac{1}{2}\epsilon\frac{\delta S}{\delta\phi^{t_s}(x)}. \tag{3.4.16}$$

Working out the changes to eq (3.4.9) and eq (3.4.11) one found that the choice

$$a = 2\sqrt{2} - 1 \tag{3.4.17}$$

$$b = 1. \tag{3.4.18}$$

makes the coefficient of order ϵ in the shift δS equal to zero.

Another possible second order integrator is the integrator scheme introduced in the appendix (A.4) (A.15) of [32]

$$\phi(x)^{t_s+\epsilon} = \phi^{t_s}(x) - \sqrt{\epsilon}\eta^{t_s}(x) - \frac{\epsilon}{3}\left(\frac{\delta S}{\delta\phi}\bigg|_{\phi^{a,t_s}(x)} + 2\frac{\delta S}{\delta\phi}\bigg|_{\phi^{b,t_s}(x)}\right), \tag{3.4.19}$$

where $\frac{\delta S}{\delta\phi}\big|_{\phi^{a,t_s}(x)}$ is the derivative of the action computed in $\phi^{a,t_s}(x)$ and

$$\phi^{a,t_s}(x) = \phi^{t_s}(x) - \frac{1}{2}\epsilon\frac{\delta S}{\delta\phi^{t_s}(x)} \tag{3.4.20}$$

$$\phi^{b,t_s}(x) = \phi^{a,t_s}(x) - \frac{3\sqrt{\epsilon}}{4}\left(\eta^{t_s}(x) + \frac{\eta^{t_s+\frac{\epsilon}{2}}(x)}{\sqrt{3}}\right). \tag{3.4.21}$$

The same analysis of before leads to a discrete Fokker-Planck equation of the form

$$\frac{1}{\epsilon}(P^{t_s+\epsilon}[\phi] - P^{t_s}[\phi]) = \frac{\delta}{\delta\phi^{t_s}(x_1)}\left(\frac{\delta S}{\delta\phi^{t_s}(x_1)} + \frac{\delta}{\delta\phi^{t_s}(x_1)}\right)P^{t_s}[\phi] + O(\epsilon^2), \tag{3.4.22}$$

which tells us that the equilibrium distribution will be $P \propto e^{-(S+O(\epsilon^2))}$.

Both (3.4.19) and (3.4.15) are second order integrators, (3.4.15) may seem at first look more efficient because it requires only two evaluation of the derivative of the action, however the extra term in the (3.4.19) permits us to minimize the coefficient in front of the ϵ^2 term [32] and have an optimal second order integrator. For this reason, in our numerical simulations, we used the integrator (3.4.19).

3.5 Langevin based Numerical Stochastic Perturbation Theory (LSPT)

Stochastic quantization based on the Langevin equation was the first method that had a numerical implementation for perturbation theory. In the literature the simulation were usually performed using the bare mass m_0 [8–10]. We find it more convenient to work with the renormalized mass m and its counterterm (q.v., (2.0.2)) such that the continuum Langevin equation reads

$$\partial_{t_s}\phi(t_s,x) = -F\big(\phi(t_s,x)\big) + \eta(t_s,x), \tag{3.5.1}$$

where $F\big(\phi(t_s,x)\big)$ denotes the functional derivative of the action (2.0.1) evaluated on the field configuration $\phi(t_s,x)$,

$$F\big(\phi(t_s,x)\big) = \frac{\delta S[\phi]}{\delta\phi(t_s,x)} = -\Delta\phi(t_s,x) + (m^2+\delta m^2)\phi(t_s,x) + \frac{g_0}{3!}\phi(t_s,x)^3. \tag{3.5.2}$$

A perturbative solution of (3.5.1) order by order in the couplings is obtained expanding the stochastic field ϕ as a power series in the couplings of the theory. In particular, in the regularized theory we can consider an expansion in terms of the bare coupling g_0,

$$\phi(x) = \sum_{k=0}^{N} \phi_k(x) g_0^k + O(g_0^{N+1}). \tag{3.5.3}$$

However for the numerical implementation of the method we find convenient to expand the field as a two-dimensional array $\phi_{k,\ell}$ of powers of g_0 and also δm^2

$$\phi(x) = \sum_{k,\ell=0}^{N} \phi_{k,\ell}(x) g_0^k (\delta m^2)^\ell + O(g_0^{N+1}). \tag{3.5.4}$$

Substituting the expansion of the stochastic field ϕ analogous to (3.5.4) into (3.5.1) gives a system of equations for the fixed order fields,

$$\begin{aligned}
\partial_{t_s} \phi^{t_s}_{0,0}(x) &= (\Delta - m^2)\phi^{t_s}_{0,0}(x) + \eta^{t_s}(x), \\
\partial_{t_s} \phi^{t_s}_{1,0}(x) &= (\Delta - m^2)\phi^{t_s}_{1,0}(x) - \frac{1}{3!}\phi^{t_s}_{0,0}(x)^3, \\
\partial_{t_s} \phi^{t_s}_{0,1}(x) &= (\Delta - m^2)\phi^{t_s}_{0,1}(x) - \phi^{t_s}_{0,0}(x),
\end{aligned} \tag{3.5.5}$$

and so on. These equations can readily be solved for the $\phi_{k,\ell}$ fields. Once a solution is obtained up to a given order in the coupling, (3.0.6) can be used to compute the perturbative expansion of any correlation function in the corresponding Euclidean field theory.

We wrote a code that integrates (3.5.1) using the scheme 3.4.19 in an automated fashion by employing order by order operations

$$(\phi \cdot \phi)(x,y) = \phi(x)\phi(y) \quad \Longrightarrow \quad (\phi \cdot \phi)_{k,\ell}(x,y) = \sum_{0\le i\le k} \sum_{0\le j\le \ell} \phi_{k-i,\ell-j}(x)\phi_{i,j}(y), \tag{3.5.6}$$

and similarly for other elementary operations. Once the solution of the Langevin equation is obtained for large enough t_s, correlation functions of the stochastic field can be expanded order by order in the couplings and evaluated stochastically by averaging over different samples of the Gaussian random fields η_i. Assuming ergodicity the average over the random field distribution on the left hand side of (3.0.6) is replaced by an average over stochastic time, and one obtains

$$\lim_{t_s\to\infty} \big\langle \mathcal{O}(t_s) \big\rangle_\eta = \langle \mathcal{O} \rangle \quad \xrightarrow{t_s = n\varepsilon} \quad \lim_{T\to\infty} \frac{1}{T} \sum_{n=0}^{T} \mathcal{O}\big(\phi(n\varepsilon)\big) = \langle \mathcal{O} \rangle + O(\varepsilon^p). \tag{3.5.7}$$

In the above relation the equivalence between correlation functions is valid order

by order in perturbation theory (q.v., (3.5.10)), whereas the power p depends on the order of the chosen integration scheme: with the integrator (3.4.19) $p = 2$ (see §3.4).[3] The perturbative expansion of generic observables of the stochastic field $\mathcal{O}(\phi)$ can be computed by iterating order by order convolution operations of the form (3.5.6), In this way one obtains the generic stochastic perturbative Observable,

$$\mathcal{O}(\phi) = \sum_{k,\ell=0}^{N} \mathcal{O}_{k,\ell}(\phi_{0,0}, \ldots, \phi_{k,\ell})\, g_0^k (\delta m^2)^\ell + O(g_0^{N+1}) \tag{3.5.8}$$

from which the perturbative expansion of the expectation value of the field $\mathcal{O}(\varphi)$ in φ^4 theory,

$$\langle \mathcal{O} \rangle = \sum_{k,\ell=0}^{N} a_{k,\ell}\, g_0^k (\delta m^2)^\ell + O(g_0^{N+1}), \tag{3.5.9}$$

is obtained up to $O(g_0^{N+1})$ corrections as,

$$\langle \mathcal{O} \rangle_\eta = \langle \mathcal{O} \rangle \quad \Longleftrightarrow \quad \langle \mathcal{O}_{k,\ell} \rangle_\eta = a_{k,\ell}. \tag{3.5.10}$$

Once the expansion (3.5.8) is known the corresponding expansion in terms of a given renormalized mass and coupling (as well as any renormalization of the field $\mathcal{O}$) is easily found (q.v., Appendix A).

We note that within the statistical uncertainties the perturbative expansions thus obtained are correct only up to systematic errors due to the discretization of the stochastic time.

The increase of the cost of LSPT with the perturbative order is dictated by the order-by-order operations necessary to integrate the discrete Langevin equation. Consequently, the computational cost of LSPT increases (roughly) with the square of the order in each coupling (q.v., (3.5.6)). However, as just mentioned, the results need to be extrapolated to zero step-size to eliminate systematic errors in the results.

[3] In practical simulations the value of T is necessarily finite, and one averages the fields only once the discrete stochastic process has equilibrated.

3.5.1 Observables in NSPT

As we saw in the previous section, the quantities $a_{k,l}$ defined in (3.5.9) are computed in NSPT using the stochastic estimators $\langle \mathcal{O}_{k,\ell}(\phi_{0,0}, \ldots, \phi_{k,\ell})\rangle_\eta$ defined in (3.5.8). After applying the renormalization procedure described in Appendix A the mass counterterms are determined and the double power series (3.5.8) and (3.5.9) can be reduce to a single power series in g

$$\mathcal{O}(\phi) = \sum_{k=0}^{N} \mathcal{O}_k(\phi_0, \ldots, \phi_k)\, g^k + O(g_0^{N+1})\,, \tag{3.5.11}$$

$$\langle \mathcal{O} \rangle = \sum_{k=0}^{N} a_k\, g^k + O(g_0^{N+1}). \tag{3.5.12}$$

The stochastic estimators $\mathcal{O}_k$ of the observables a_k have variance equal to

$$\mathrm{Var}(\mathcal{O}_k) = \langle \mathcal{O}_k^2 \rangle_\eta - \langle \mathcal{O}_k \rangle_\eta^2. \tag{3.5.13}$$

The expansion of the variance of $\mathcal{O}$ in g is

$$\mathrm{Var}(\mathcal{O}) = \langle \mathcal{O}^2 \rangle - \langle \mathcal{O} \rangle^2 = \sum_{k=0}^{N} b_k g^k. \tag{3.5.14}$$

Beyond the tree level in general the variance of the perturbative coefficients of an observable (3.5.13) differs from the coefficient of the perturbative expansion of the variance (3.5.14) of the same observable, i.e.

$$\mathrm{Var}(\mathcal{O}_k) \neq b_k \qquad \forall k > 0. \tag{3.5.15}$$

The quantities b_k correspond to a correlation functions in the static theory while the $\mathrm{Var}(\mathcal{O}_k)$ for $k > 0$ do not, further their value depends on the detail of the NSPT algorithm. This is different from a non perturbative simulation where the variance of an observable is determined by the theory and thus independent on the algorithm used.

In addition, as the fields entering in the average in (3.5.7) are generated by a

Markov process, successive field configurations are correlated. The autocorrelation function of an observable $\mathcal{O}_k$ was defined in (B.1.5) and in this case it takes the form

$$\Gamma_{\mathcal{O}_k}(t_s) = \langle(\mathcal{O}_k(t_s) - a_k)(\mathcal{O}_k(0) - a_k)\rangle_\eta. \tag{3.5.16}$$

As in the case of the variance it differs from the perturbative coefficient in the expansion of the autocorrelation $b_k(t)$ of the same observable

$$\Gamma_{\mathcal{O}}(t_s) = \langle(\mathcal{O}(t_s) - \langle\mathcal{O}\rangle)(\mathcal{O}(0) - \langle\mathcal{O}\rangle)\rangle_\eta = \sum_{k=0}^{N} b_k(t_s) g^k \tag{3.5.17}$$

beyond the tree level, i.e., $\Gamma_{\mathcal{O}_k}(t_s) \neq b_k(t_s)$ in general for $k > 0$.

3.5.2 Variances and Autocorrelations in LSPT

Here we show how the theoretical discussion of the previous section on the renormalization of the Langevin equation can be used to investigate the variance and the autocorrelation of the estimator of $\mathcal{O}_k$. The variance $\mathrm{Var}(\mathcal{O}_k)$ (3.5.13) can be computed in the $d+1$ dimensional theory as

$$\mathrm{Var}(\mathcal{O}_k) = \int_{\omega,\nu} \langle\mathcal{O}_k(\omega)\mathcal{O}_k(\nu)\rangle_\eta - \langle\mathcal{O}_k(\omega)\rangle_\eta\langle\mathcal{O}_k(\nu)\rangle_\eta. \tag{3.5.18}$$

In general this will contain graphs up to order $2k$ in the correlator $\left\langle\mathcal{O}_k^2(\omega)\right\rangle_\eta$. If these graphs have divergences they will not be cancelled by the renormalization procedure since in $\mathcal{O}_k$ there are only counterterms up to order k. In the Langevin $d+1$ dimensional theory the degree of divergence of a graphs (§3.3.7) is

$$deg\,\mathcal{G} = 3 - \frac{3}{2}E + E_G\,, \tag{3.5.19}$$

The maximum degree of divergence is when $E_G = E - 1$ which means that the 1PI graph is generated by the contraction of the noise within the same tree graphs of the field in figure 3.1 or, more precisely an E-point 1PI correlation function at

order k is a sum of terms like

$$\mathcal{O}_k = \sum_{i_1,\cdots i_E} \langle \phi_{i_1} \cdots \phi_{i_E} \rangle_\eta^{1PI} \delta_{i_1+\cdots+i_E,k} \tag{3.5.20}$$

$$= (\langle \phi_k \phi_0 \cdots \phi_0 \rangle_\eta^{1PI} + \langle \phi_{k-1} \phi_1 \cdots \phi_0 \rangle_\eta^{1PI} + \langle \phi_{k-2} \phi_1 \phi_1 \cdots \phi_0 \rangle_\eta^{1PI} + \langle \phi_{k-2} \phi_2 \cdots \phi_0 \rangle_\eta^{1PI} \cdots) + (\text{permutations}). \tag{3.5.21}$$

The first term in the second line will have $E_G = E-1$, the second $E_G = E-2$, the third $E_G = E-3$, and the fourth $E_G = E-2$. In the correlator $\langle \mathcal{O}_k(\omega) \mathcal{O}_k(\nu) \rangle_\eta$ 1PI subgraphs of order larger than k will have $E_G < E-1$, because in $\mathcal{O}_k$ there is no field ϕ_s of order $s > k$. Thus the maximum degree of divergence of the graphs of $\langle \mathcal{O}_k(\omega) \mathcal{O}_k(\nu) \rangle_\eta$ will occur when $E_G = E-2$

$$\max \left(deg\, \mathcal{G}_{\langle \mathcal{O}_k(\omega) \mathcal{O}_k(\nu) \rangle_\eta} \right) = 1 - \frac{1}{2} E\,, \tag{3.5.22}$$

which means that only a two-point subgraph can have $deg\mathcal{G} = 0$ and thus be logarithmically divergent, all the other must be convergent. Therefore the correlator $\left\langle \mathcal{O}_k^2(\omega) \right\rangle_\eta$ will be at the most logarithmically divergent.

This feature was already shown by Martin Lüscher [33] for a different point of view, writing the dynamical Langevin $d+1$ theory as a path integral for all order of the stochastic field ϕ_k.

The absence of power divergences in the variance of the perturbative coefficient is a remarkable property of the Langevin equation; we observed that in ISPT (§5) and GHMD (§4) when the parameters are tuned to minimise the autocorrelations the variances grow as powers of the lattice spacing as the continuum limit is approached (§6).

The same argument can be applied to the autocorrelation function

$$\Gamma_{\mathcal{O}_k}(t_s - t'_s) = \int_{\omega,\nu} e^{-it_s\omega - it'_s\nu} \langle (\mathcal{O}_k(\omega) - a_k)\, (\mathcal{O}_k(\nu) - a_k) \rangle \tag{3.5.23}$$

concluding that after the renormalization of Langevin equation it is at the most logarithmically divergent.

However in a simulation the renormalization factor for the Langevin time is not included so the autocorrelation function will have a divergence proportional to ω which for dimensional reasons will grow $1/a^2$ as the continuum limit is approached.

Chapter 4

NSPT based on GHMD algorithms

The idea of stochastic perturbation theory is not limited to the Langevin equation. Any stochastic differential equation which satisfies a property analogous to (3.0.6) can provide a means of performing stochastic perturbation theory. One interesting example is given by the stochastic molecular dynamics (SMD) equations (4.3.1). In a non-perturbative context these were first considered in [34], and were recently studied in detail in [26]. Similarly, one can set up perturbation theory in terms of the Hybrid Molecular Dynamics (HMD) equations [35]. This observation suggests the possibility of defining NSPT based on the discretization of these SDEs or of ergodic variants of the molecular dynamics (MD) equations; such as the Kramers [36–39] and HMD algorithm respectively [40]. Experience with conventional non-perturbative lattice field theory simulations would suggest the advantages of reformulating NSPT in terms of these algorithms rather than Langevin-based ones. However determining their efficiency in this context, in particular their continuum scaling, is not a trivial issue. The results for the free field theory [41] provide a complete understanding of the lowest perturbative order dynamics.

On the other hand the analysis in §3.3 can not be applied in these algorithms because the dynamical $d+1$ dimensional theory defined form the SMD and HMD equations is not renormalizable, a direct calculation in [26] shows that the four-point function contains a non-local singularity that cannot be canceled by

including local counterterms in the stochastic SMD and HMD equations. This feature, in general, precludes analytic control over the continuum scaling of these algorithms in the interacting theory. In the case of NSPT this means a lack of control of the behaviour of the higher-order fields. Consequently, the efficiency of these algorithms in the context of NSPT must be addressed numerically; in particular the situation could be substantially different from both the free case and the case where the full theory is simulated.

In the following we introduce the HMD and Kramers algorithms (see [41] and references therein for their definition) in a full non-perturbative context and then we will describe the required modifications for their NSPT formulations. These NSPT implementations are all inexact algorithms, as we do not know how to add a Metropolis step that would be valid for arbitrary values of the coupling beyond leading (free field) order. These algorithms will be called HSPT and KSPT, respectively, they are a particular case of the more general Generalized Hybrid Molecular Dynamics (GHMD) algorithm [41]. In chapter 4.4 we will give a proof of the convergence for the NSPT formulations of the GHMD algorithm, however we limited the numerical study to the HMD and Kramers which, based on both the expectations from free field theory and from non-perturbative lattice field theory, appear to be natural sub-classes of the GHMD algorithm to consider.

4.1 Hybrid Algorithms

The Hybrid Molecular Dynamics (HMD) [35], is an algorithm for simulation of quantum field theory. We will describe it in the case of a ϕ^4 theory defined from the action 2.0.1. The algorithm consists in the iteration of the following steps:

1. Generate "fictitious" momenta π with gaussian distribution

$$P[\pi] \sim \frac{1}{Z} e^{\frac{1}{2}\sum_x \pi^2(x)}; \tag{4.1.1}$$

2. Introduce the stochastic time t_s [1] with the identification

$$\pi(t_s = 0, x) = \pi(x) \quad \phi(t_s = 0, x) = \phi(x) \tag{4.1.2}$$

and integrate of the Hamilton equation in t_s up to a trajectory τ

$$\frac{d\pi(t_s, x)}{dt_s} = -\frac{\partial H[\pi, \phi]}{\partial \phi(t_s, x)} \qquad \frac{d\phi(t_s, x)}{dt_s} = \frac{\partial H[\pi, \phi]}{\partial \pi(t_s, x)} \tag{4.1.3}$$

where the Hamiltonian function H is defined as

$$H[\pi, \phi] = \frac{1}{2}\sum_x \pi^2(t_s, x) + S[\phi(t_s, x)] \tag{4.1.4}$$

and S is the lattice action 2.0.1;

The HMD is a Markov process. A Markov process is a stochastic procedure which generates a new configuration ϕ' from its predecessor ϕ with probability $P_M(\phi \to \phi')$. Any Markov process on a compact space will converge to a unique fixed point distribution P_s provided that it is ergodic and that it satisfies detailed balance

$$P_s(\phi) P_M(\phi \to \phi') = P_s(\phi') P_M(\phi' \to \phi). \tag{4.1.5}$$

in the following we will show that $P_s(\phi) \propto e^{-S[\phi]}$ satisfies the detailed balance

[1] This stochastic time is different from the one introduced in the case of the Langevin, in particular its dimension is different

4.1.5 with the transition probability of the HMD P_M. After one iteration of the HMD the probability P_H to end in a configuration (π', ϕ') starting from (π, ϕ) is

$$P_H((\pi, \phi) \to (\pi', \phi')) = \delta[(\pi', \phi') - (\pi(\tau), \phi(\tau))], \tag{4.1.6}$$

Where $(\pi(\tau), \phi(\tau))$ is the result of the Hamilton dynamics up to a trajectory of τ. The marginal transition probability for the field ϕ alone is

$$P_M(\phi \to \phi') = \int D\pi D\pi' e^{-\frac{1}{2}\sum_x \pi^2(x)} \delta[(\pi', \phi') - (\pi(\tau), \phi(\tau))]. \tag{4.1.7}$$

The Hamilton dynamics preserves the Hamiltonian i.e., $H[\pi(0), \phi(0)] = H[\pi(\tau), \phi(\tau)]$, and it is reversible

$$P_H((\pi, \phi) \to (\pi', \phi')) = P_H((-\pi', \phi') \to (-\pi, \phi)) = \delta[(-\pi, \phi) - (-\pi'(\tau), \phi'(\tau))]. \tag{4.1.8}$$

Using the fact that $P_s(\phi) e^{-\frac{1}{2}\sum_x \pi^2(x)} = e^{-H[\pi,\phi]}$ we have

$$P_s(\phi) P_M(\phi \to \phi') = \int D\pi D\pi' P_S(\phi) e^{-\frac{1}{2}\sum_x \pi^2(x)} \delta[(\pi', \phi') - (\pi(\tau), \phi(\tau))] \tag{4.1.9}$$

$$= \int D\pi D\pi' e^{-H[\pi,\phi]} \delta[(-\pi, \phi) - (-\pi'(\tau), \phi'(\tau))] \tag{4.1.10}$$

$$= \int D\pi D\pi' e^{-H[\pi'(\tau),\phi'(\tau)]} \delta[(-\pi, \phi) - (-\pi'(\tau), \phi'(\tau))] \tag{4.1.11}$$

$$= \int D\pi D\pi' e^{-H[\pi',\phi']} \delta[(\pi, \phi) - (\pi'(\tau), \phi'(\tau))] = P_s(\phi') P_M(\phi' \to \phi), \tag{4.1.12}$$

where from the first to the second line we used the (4.1.8), from the second to the third line we use the properties of the delta function, from the third to the fourth we use the fact that the Hamiltonian dynamics preserves the Hamiltonian and we change integration variable $\pi \to -\pi$ using the fact that $H[\pi, \phi] = H[-\pi, \phi]$.

If the dynamics in t_s defined in the step two is only approximated as the case of a numerical integration of the Hamilton's equations then the algorithm it is not exact. A way to correct this is to include a Metropolis accept/reject step, so the

new configuration generated after one iteration of HMC will be

$$(\pi', \phi') = \begin{cases} (\pi(\tau), \phi(\tau)), & \text{with probability} \quad \min(1, e^{-\Delta H}), \\ (\pi, \phi), & \text{otherwise}, \end{cases} \tag{4.1.13}$$

where $\Delta H = H[\pi', \phi'] - H[\pi, \phi]$. The resulting algorithm is called Hybrid Monte Carlo (HMC) [40] and consist in:

1. Generation of "fictitious" momenta π with gaussian distribution;
2. An approximate integration of the Hamilton equation which is area preserving and reversible 4.1.8;
3. A flip of the momenta $F : \pi \to -\pi$;
4. An accept/reject (4.1.13).

With similar steps to (4.1.9) It is possible to show that the HMC satisfies the detailed balance [40].

4.1.1 Inexact algorithm

If the dynamics used in the second step of the HMD algorithm is replaced by an approximate integrator we obtain an inexact algorithm. Since we are interested in the NSPT applications of the HMD algorithm it is useful study the inexact algorithm, we do not know how to add an accept/reject step to make the algorithm exact. In this section we will discuss the technique and main results obtained in [42] on the inexact HMD algorithm. We will discuss the case where a symplectic integrator is used. For a time-independent Hamiltonian the time evolution operator of a symplectic integrator can be written as

$$\exp\left(\tau \tfrac{d}{dt_s}\right) \equiv \exp\left(\tau \left\{ \frac{\partial \pi(t_s, x)}{\partial t_s} \frac{\partial}{\partial \pi(t_s, x)} + \frac{\partial \phi(t_s, x)}{\partial t_s} \frac{\partial}{\partial \phi(t_s, x)} \right\}\right) \tag{4.1.14}$$

$$= \exp\left(\tau \left\{ -\frac{\partial H[\pi, \phi]}{\partial \phi(t_s, x)} \frac{\partial}{\partial \pi(t_s, x)} + \frac{\partial H[\pi, \phi]}{\partial \pi(t_s, x)} \frac{\partial}{\partial \phi(t_s, x)} \right\}\right) \equiv e^{\tau \widehat{H}} \tag{4.1.15}$$

where $\widehat{H}$ is the Hamiltonian vector field and we made use of Hamilton's equations. To make the notation lighter we will drop all the dependencies of the field upon stochastic time t_s and space x, so the last equation become

$$e^{\tau \widehat{H}} \equiv \exp\left(\tau \left\{-\partial_\phi H \partial_\pi + \partial_\pi H \partial_\phi\right\}\right). \tag{4.1.16}$$

The Hamiltonian is a sum of a kinetic energy $T(\pi)$ which depend only on π and the potential energy $S(\phi)$ which is a function of the field ϕ only. Thus we can define the operators

$$\Phi \equiv (\partial_\pi T)\partial_\phi \quad \text{and} \quad \Pi \equiv -(\partial_\phi S)\partial_\pi. \tag{4.1.17}$$

Using the Taylor series we have

$$e^{\tau\Phi} f[\pi, \phi] = f[\pi, \phi + \partial_\pi H] \tag{4.1.18}$$

$$e^{\tau\Pi} f[\pi, \phi] = f[\pi - \partial_\phi H, \phi]. \tag{4.1.19}$$

The leapfrog scheme is a symmetric symplectic integrator that can be written as

$$U(\tau) = U_0(\delta\tau)^{\tau/\delta\tau} = \left(e^{\delta\tau\Pi/2} e^{\delta\tau\Phi} e^{\delta\tau\Pi/2}\right)^{\tau/\delta\tau}. \tag{4.1.20}$$

Using the Baker-Campbell-Hausdorff (BCH) formula we get:

$$U(\tau) = \left(\exp\left[(\Pi+\Phi)\delta\tau - \frac{1}{24}\big([\Pi,[\Pi,\Phi]] + 2[\Phi,[\Pi,\Phi]]\big)\delta\tau^3 + O(\delta\tau^5)\right]\right)^{\tau/\delta\tau} \tag{4.1.21}$$

$$= \exp\left[(\Pi+\Phi)\tau - \frac{1}{24}\big([\Pi,[\Pi,\Phi]] + 2[\Phi,[\Pi,\Phi]]\big)\tau\delta\tau^2 + O(\delta\tau^4)\right] \equiv e^{\tau\widehat{H}'}, \tag{4.1.22}$$

where $[\widehat{A}, \widehat{B}] = \widehat{A}\widehat{B} - \widehat{B}\widehat{A}$ is the commutator. The BCH formula tells us that for the leapfrog integrator the trajectories through phase space are integral curves of the vector field

$$\widehat{H}' = (\Pi+\Phi) - \frac{1}{24}\big([\Pi,[\Pi,\Phi]] + 2[\Phi,[\Pi,\Phi]]\big)\delta\tau^2 + O(\delta\tau^4). \tag{4.1.23}$$

This means that the Hamiltonian H' such that

$$\widehat{H}' = -(\partial_\phi H')\partial_\pi + (\partial_\pi H')\partial_\phi \tag{4.1.24}$$

is exactly conserved. The Hamiltonian H' is called shadow Hamiltonian. H' is obtained using the identity proved in Appendix A of [42]

$$[\widehat{A}, \widehat{B}] = \widehat{\{A, B\}} \tag{4.1.25}$$

where

$$\{A, B\} = (\partial_\pi A)\partial_\phi B - (\partial_\phi A)\partial_\pi B \tag{4.1.26}$$

is the Poisson bracket. The (4.1.25) tells us that H' is obtained substituting the commutator with the Poisson bracket in the expression for $\widehat{H}'$, thus for the leapfrog integration scheme the shadow Hamiltonian is

$$\begin{aligned} H' &= H - \frac{1}{24}\left(\{S, \{S, T\}\} - 2\{T, \{S, T\}\}\right)\delta\tau^2 + O(\delta\tau^4) \\ &= H - \frac{1}{24}\left(-2(\partial_\phi S)^2 + 2\pi^2\partial_\phi^2 S\right)\delta\tau^2 + O(\delta\tau^4). \end{aligned} \tag{4.1.27}$$

or with the full notation

$$H' = H - \frac{1}{24}\sum_x \left(-2\left(\frac{\delta S}{\delta\phi(x)}\right)^2 + 2\pi^2(x)\frac{\delta^2 S}{\delta\phi^2(x)}\right)\delta\tau^2 + O(\delta\tau^4). \tag{4.1.28}$$

In general we do not expect that the inexact HMD will have either e^{-H} or $e^{-H'}$ as its equilibrium distribution. We will parametrize the equilibrium distribution as $e^{-(S+\Delta S)}$ where e^{-S} is the desired one and ΔS a shift. The condition to be a fixed point is

$$e^{-(S[\phi']+\Delta S[\phi'])} = \int D\pi D\phi e^{-(S[\phi]+\Delta S[\phi])} e^{\frac{1}{2}\pi^2}\delta(\phi' - \phi'') \tag{4.1.29}$$

$$= \int D\pi D\phi e^{-(H[\pi,\phi]+\Delta S[\phi])}\delta(\phi' - \phi''). \tag{4.1.30}$$

where $(\pi'', \phi'') = U(\tau)(\pi, \phi)$ is the phase space configuration reached at the end of

a trajectory of length τ. We change the integration variable to (π'', ϕ'') obtaining

$$e^{-(S[\phi']+\Delta S[\phi'])} = \int D\pi'' D\phi'' (\det U_*)^{-1} e^{-(H+\Delta S)\circ U^{-1}[\pi'',\phi'']} \delta(\phi' - \phi''). \quad (4.1.31)$$

If we are using an area-preserving integrator then the Jacobian $\det U_* = \det \frac{\partial(\pi'',\phi'')}{\partial(\pi,\phi)} = 1$, and if it is reversible $U^{-1} = F \circ U \circ F$, with $F : \pi \to -\pi$. Both condition are fulfilled by the leapfrog integrator. Using the reversibility and the fact that the Hamiltonian in invariant under momentum flip

$$\begin{aligned}(H + \Delta S) \circ U^{-1}[\pi'', \phi''] &= (H + \Delta S) \circ F \circ U \circ F[\pi'', \phi''] = (H + \Delta S) \circ U[-\pi'', \phi''] \\ &= (H + \Delta S) \circ U[-\pi'', \phi''] - (H + \Delta S)[-\pi'', \phi''] + (H + \Delta S)[-\pi'', \phi''] \\ &= (H + \Delta S)[\pi'', \phi''] + \delta(H + \Delta S), \end{aligned} \quad (4.1.32)$$

where $\delta(H+\Delta S)$ is the change of the Hamiltonian plus the shift in the distribution over a trajectory with an momentum flip $\delta(H + \Delta S) = (H + \Delta S) \circ U[-\pi'', \phi''] - (H + \Delta S) \circ U[-\pi'', \phi'']$. With the use of the (4.1.32) the (4.1.31)

$$e^{-(S[\phi']+\Delta S[\phi'])} = \int D\pi'' D\phi'' e^{-(H+\Delta S)[\pi'',\phi'']-\delta(H+\Delta S)} \delta(\phi' - \phi'') \quad (4.1.33)$$

$$= \int D\pi'' e^{-(H+\Delta S)[\pi'',\phi']-\delta(H+\Delta S)} \quad (4.1.34)$$

$$= e^{-(S[\phi']+\Delta S[\phi'])} \int D\pi'' e^{-\frac{1}{2}\pi''^2} e^{-\delta(H+\Delta S)}. \quad (4.1.35)$$

Thus the condition for the fixed point distribution for a reversible area preserving integrator is

$$\int D\pi'' e^{-\frac{1}{2}\pi''^2} e^{-\delta(H+\Delta S)} = 1. \quad (4.1.36)$$

For the HMD algorithm we cannot derive an explicit expression for ΔS, however from the (4.1.27) we can see that H' differs from H by terms of $O(\delta\tau^2)$. This implies that $\delta H = O(\delta\tau^2)$ and thus from the (4.1.36) we deduce that $\delta\Delta S = O(\delta\tau^2)$, so ΔS in general will be of $O(\delta\tau^2)$.

In general a symplectic area preserving reversible integrator of order n will conserve a shadow Hamiltonian H' which will differs from H by terms $O(\delta\tau^n)$ and similar consideration will lead to $\Delta S = O(\delta\tau^n)$.

4.1.2 Free theory case

The HMD algorithm for the free case has an extra complication, the fact that for some values of the trajectory length it is not ergodic [41, 43];. This can be seen explicitly in Fourier space. The hamiltonian for the free theory in Fourier space (2.3.2) is

$$H[\pi,\phi] = \frac{1}{2}\sum_{p\in\tilde{\Omega}} \big(\pi(-p)\pi(p) + \omega_p^2\phi(-p)\phi(p)\big) \tag{4.1.37}$$

where $\tilde{\Omega}$ defined in the (2.0.6) and the frequencies are

$$\omega_p^2 \equiv m^2 + 4\sum_{\mu=1}^{4} \sin^2\frac{\pi p_\mu}{L}. \tag{4.1.38}$$

One step of leapfrog (4.1.20) from the initial time t_0 to $t_0+\delta t$ can be written as

$$\begin{pmatrix}\phi^{t_0+\delta t}(p)\\ \pi^{t_0+\delta t}(p)\end{pmatrix} = U(\delta t)\begin{pmatrix}\phi^{t_0}(p)\\ \pi^{t_0}(p)\end{pmatrix}. \tag{4.1.39}$$

with

$$U(\delta t) = \begin{pmatrix} 1-\frac{1}{2}(\omega_p\delta t)^2 & \delta t \\ -\omega_p^2\delta t + \frac{1}{4}\omega_p^4\delta t^3 & 1-\frac{1}{2}(\omega_p\delta t)^2\end{pmatrix}. \tag{4.1.40}$$

The fact that the diagonal element contain only even powers of δt and the off diagonal only odd powers suggests the following parametrization

$$U(\delta t) = \begin{pmatrix} \cos[\kappa(\delta t)\omega_p\delta t] & \frac{\sin[\kappa(\delta t)\omega_p\delta t]}{\rho(\kappa(\delta t))} \\ -\rho(\delta t)\sin[\kappa(\delta t)\omega_p\delta t] & \cos[\kappa(\delta t)\omega_p\delta t]\end{pmatrix}. \tag{4.1.41}$$

The latter parametrization holds for an arbitrary symplectic integrator, in § 4.4 we will treat the generic case. Enforcing equality between 4.1.41 and 4.1.40 we can compute the coefficients of the Taylor expansion of the functions

$$\kappa(\delta t) = 1 + \frac{1}{24}(\omega_p\delta t)^2 + \frac{3}{640}(\omega_p\delta t)^4 + O(\delta t^6). \tag{4.1.42}$$

$$\rho(\delta t) = \omega_p\left(1 - \frac{1}{8}(\omega_p\delta t)^2 - \frac{1}{128}(\omega_p\delta t)^4\right) + O(\delta t^6). \tag{4.1.43}$$

The parametrization (4.1.41) make it easier to compute the operator $U(\tau) = U(N\delta t) = U(\delta t)^N$ for arbitrary long trajectory τ. We assume the induction hypothesis

$$U(\tau) = \begin{pmatrix} \cos[\kappa(\delta t)\omega_p\tau] & \frac{\sin[\kappa(\delta t)\omega_p\tau]}{\rho(\kappa(\delta t))} \\ -\rho(\delta t)\sin[\kappa(\delta t)\tau] & \cos[\kappa(\delta t)\omega_p\tau] \end{pmatrix} \tag{4.1.44}$$

from this follows that $U(\tau)U(\delta t) = U(\tau+\delta t)$ using trigonometric identities. After m trajectory of HMD the field will be

$$\phi^m(p) = (\cos[\kappa(\delta t)\omega_p\tau])^m\phi^0(p) + \frac{\sin[\kappa(\delta t)\omega_p\tau]}{\rho(\kappa(\delta t))}\sum_{u=0}^{m-1}(\cos[\kappa(\delta t)\omega_p\tau])^{m-1-u}\pi^u(p), \tag{4.1.45}$$

Thus the autocorrelation function is

$$\langle\phi^{m+l}(p)\phi^m(p)\rangle = \cos^l[\kappa(\delta t)\omega_p\tau]\langle\phi^m(p)\phi^m(p)\rangle. \tag{4.1.46}$$

If $\tau = 2\pi n/\omega_p\kappa(\delta t)$ the mode with frequency ω_p never decorrelates. Increasing L/a the density of Fourier modes increase and almost every choice of τ will be close to a multiple of $\tau = 2\pi n/(\omega_p\kappa(\delta t))$. The solution to this problem is simple and is to randomize the trajectory length τ [43].

4.2 NSPT based on HMD algorithms (HSPT)

In this section we will set up perturbation theory in terms of the Hybrid Molecular Dynamics (HMD), we will refer to the the resulting algorithm as HSPT. In the case of the HMD algorithm, the basic field evolution is described by the MD equations,

$$\partial_{t_s}\phi(t_s,x) = \pi(t_s,x), \qquad \partial_{t_s}\pi(t_s,x) = -F\big(\phi(t_s,x)\big), \tag{4.2.1}$$

where $F\big(\phi(t_s,x)\big)$ is given by (3.5.2), and π is the momentum field conjugate to ϕ. Similarly to the Langevin case (cf. §3.5), in the context of NSPT both fields ϕ and π are assumed to have an expansion of the form (3.5.4). All operations in the following are thus intended to be performed in an order-by-order fashion

(q.v., (3.5.6)). The momentum field π is sampled from a Gaussian distribution with zero mean and unit variance at the beginning of each trajectory ($t_s = t_0$); the refreshed momentum initially only has a non-zero lowest order component. In formulas

$$\left\langle \pi_{0,0}(t_0, x) \right\rangle_\pi = 0, \qquad \left\langle \pi_{0,0}(t_0, x)\pi_{0,0}(t_0, y) \right\rangle_\pi = \delta_{xy}, \tag{4.2.2}$$

and $\pi_{k,\ell}(t_0, x) = 0$ if either $k > 0$ or $\ell > 0$, where $\langle \cdots \rangle_\pi$ denotes the average over the momentum field distribution at the beginning of a trajectory. The momentum field will acquire higher-order components during the MD evolution (4.2.1) from the time t_0 at which it was refreshed to time $t_s = t_0 + \tau$, where τ is the trajectory length. Numerically the MD evolution is determined by discretizing the simulation time as $t_s = n\delta t$, with $n \in \mathbb{N}$ and δt the step-size, and employing a suitable integration scheme. Expectation values of generic observables are then obtained similarly to (3.5.7) by averaging over sequences of trajectories. For the numerical integration of the MD equations it is convenient to rely on some symplectic integration scheme. Using them we will prove the convergence in §4.4. Symplectic integrators can systematically be improved, and sophisticated symplectic integrators are readily available (q.v., [44] for a discussion). Moreover, once an efficient symplectic integrator is found for a scalar theory, it can be extended to non-Abelian theories in a straightforward manner. For the numerical simulation we used the fourth order integrator defined by equations (63) and (71) of [45], which we refer to as OMF4 integrator. In [46] it was proved that this integrator has $O(\delta t^4)$ errors in the distribution of the free theory for a trajectory of one step, in §4.4 we prove that the equilibrium probability distribution does not change if a trajectory of many steps is taken. The arguments of §4.1.1 and [42] shows that in a non-perturbative simulation the equilibrium distribution is also expected to differ from the desired one by terms of $O(\delta t^4)$ for small enough δt, this is also expected to be true order-by-order in perturbation theory. More precisely, we expect in general that the equilibrium probability distribution of fields generated through the HMD algorithm with some symplectic integrator of order p is, for small enough step-size δt, of the form $\bar{P}(\phi) \propto e^{-\bar{S}(\phi)}$, where $\bar{S} = S + \Delta S$ with $\Delta S = O(\delta t^p)$ (see [42] for more details). Consequently, since $\Delta S \propto V$, one may argue that in order to keep the step-size errors in the equilibrium distribution (approximately) constant as the system size V is

increased, one needs to keep the quantity $y \equiv V\delta t^p$ fixed. It is clear that this is feasible only if efficient high-order integrators are employed. We could include an accept/reject step in the HMD evolution of the lowest order field $\phi_{0,0}$. In this case the equilibrium probability distribution would be correct at this order. Keeping the acceptance probability fixed in this case would then require $x = V\delta t^{2p}$ to be fixed [41], which is a less stringent condition than keeping y fixed. However, it is not clear what the step-size errors would be for the higher-order components of the field in this case. We note that although keeping y fixed would keep systematic errors in generic correlators approximately constant as the system size is increased, this is probably an over-conservative condition if one is interested in (connected) correlation functions of local fields [31, 47].

The problem that the evolution of the lowest-order (free) field $\phi_{0,0}$ is not ergodic §4.1.2 affects the evolution of the higher-orders of the Stochastic Perturbation Theory so we need to randomize the trajectory length τ in this case [43]. The choice of distribution for the trajectories lengths may affect the autocorrelations of the Markov chain. In our implementation we fixed the step-size δt, while choosing the number of steps n composing the trajectory according to a binomial distribution with mean $\langle n \rangle$. This defines the average trajectory length to be $\langle \tau \rangle = \langle n \rangle \delta t$.

As shown in [26] the analysis of §3.3 can not be applied. Hence apart from a few cases we do not know the scaling of the variances and the autocorrelations of the NSPT observables defined in §3.5.1. Thus we will investigate numerically the algorithm.

We conclude by pointing out that if one chooses $\tau = \delta t$, i.e., the trajectory consists of a single step, then the HMD algorithm effectively integrates the Langevin equation (3.0.2) (q.v., [42] and below). In other words, in this case the algorithm just described can be interpreted as a particular integration scheme for the Langevin equation.

4.3 KSPT

Having defined HSPT in terms of the HMD algorithm, another interesting possibility to consider is NSPT based on Kramers algorithm (also called stochastic molecular dynamics). This algorithm was proposed long ago in the context of non-perturbative field theory simulations by Horowitz [36, 37], and recently reconsidered in [39]. In this case, the stochastic equations governing the field dynamics are given by

$$\partial_{t_s}\phi(t_s,x) = \pi(t_s,x), \qquad \partial_{t_s}\pi(t_s,x) = -\gamma\pi(t_s,x) - F\big(\phi(t_s,x)\big) + \eta(t_s,x). \tag{4.3.1}$$

Here $F\big(\phi(t_s,x)\big)$ is still defined by (3.5.2), while $\eta(t_s,x)$ is a Gaussian random field satisfying

$$\big\langle \eta(t_s,x)\big\rangle_\eta = 0, \qquad \big\langle \eta(t_s,x)\eta(t'_s,y)\big\rangle_\eta = 2\gamma\delta(t_s-t'_s)\delta_{xy}, \tag{4.3.2}$$

where $\gamma > 0$ is a free parameter (see below). We observe that as special cases the (non-ergodic) MD equations (4.2.1) are obtained when $\gamma = 0$ while, up to a rescaling of stochastic time, the Langevin equation (3.0.2) is obtained for $\gamma \to \infty$ (q.v., [26]).

The implementation of Kramers algorithm is as follows. Starting from some arbitrary initial values for the fields $\phi(0,x)$ and $\pi(0,x)$, the MD equations corresponding to (4.3.1) with $\gamma = 0$ are integrated from $t_s = 0$ to $t'_s = \delta t$ through a single step of a given numerical integration scheme. The value of δt thus defines the step-size of the integrator. After this MD step, the effect of the γ term and the coupling to the random field η is taken into account by *partially* refreshing the momentum field: the momentum field $\pi(t'_s,x)$ is replaced by

$$\pi'(t'_s,x) = e^{-\gamma\delta t}\pi(t'_s,x) + \sqrt{1-e^{-2\gamma\delta t}}\,\eta(t'_s,x), \tag{4.3.3}$$

where the noise field is here normalized such that

$$\big\langle \eta(n\delta t,x)\eta(n'\delta t,y)\big\rangle_\eta = \delta_{nn'}\delta_{xy}. \tag{4.3.4}$$

These elementary steps are then alternated, and expectation values of generic observables of the field are obtained as in (3.5.7) by averaging over a long Monte Carlo history, after they have reached equilibrium.

In a KSPT implementation, the fields ϕ and π are assumed to have an expansion of the form (3.5.4), and just as in the Langevin case the random field η only has a lowest-order component. Hence, during the partial refreshment (4.3.3) only the lowest-order component of the momentum field $\pi_{0,0}$ is affected by the random field η, while the higher-order components are just rescaled by the factor $e^{-\gamma\delta t}$. In the case where $\gamma \to \infty$ (the Langevin limit) the algorithm described is just the single step HSPT algorithm.

Having defined the algorithm, some comments are in order. First of all, as shown by Horowitz's analysis [36], the partial momentum refreshment (4.3.3) integrates exactly the corresponding terms in (4.3.1). Further in [36] it was shown that in a non perturbative simulation the error in the distribution is given by the integration of the MD equations in discrete steps, this result is expected to be valid order by order in pertubation theory. In §4.4 we will give an explicit formula for the distribution probability of the lowest order and we will prove the convergence of the algorithm order by order in perturbation theory. For the present work, we employed the very same OMF4 integrator used we used for HSPT: we therefore expect $O(\delta t^4)$ step-size errors.

Secondly, we notice that in an exact implementation, i.e., when a Metropolis accept/reject step is included, free field theory analysis of [41] show that KSPT is performing worse than HSPT, at least close to the continuum limit. Including a Metropolis accept/reject step leads to the critical exponent for the *cost* of the algorithms being $z = 1$ for HMC but $z = 3/2$ for the exact Kramers algorithm (KMC). However, in the case of NSPT one is limited to *inexact* algorithms, so the computations have to be performed in a parameter regime where the effect of step-size errors on expectation values are smaller than some specified statistical accuracy, as otherwise some extrapolation in the step-size would be necessary. In this regime, corresponding to the case where the Metropolis acceptance probability

would be close to one, the two algorithms have comparable performances [41].[2]

Similar to HSPT the technique used in §3.3 to analyse the LSPT algorithm can not be used due to the lack of renormalizability shown in [26].

KSPT is also interesting due to the following property. As mentioned before, the SMD equations (4.3.1) approach the Langevin equation (3.0.2) in the limit $\gamma \to \infty$. In lattice field theory, this limit can be taken simultaneously with the continuum limit if γ is kept fixed in lattice units while $a \to 0$ [26]. In this limit the algorithm described above integrates the Langevin equation as the continuum limit of the theory is approached. Consequently, the considerations on the continuum scaling of the LSPT algorithms discussed in §3.5 directly apply to KSPT at fixed γ. Although the scaling of these algorithms is expected to be the same, in the case of KSPT the parameter γ may be fixed to some finite value for which the algorithm may be more efficient. This will be addressed in detail in §6.0.5.

[2]It is worth pointing out that even in the exact case, the critical exponent for all GHMC algorithms in the free case can be improved by using higher-order integrators for the MD equations.

4.4 Convergence NSPT based on GHMD algorithms

Both HMD and Kramer algorithm introduced in the previous sections are special case of the Generalised Hybrid Molecular Dynamics (GHMD). The (GHMD) cosist in:

- An approximate integration, area preserving and reversible, of Hamilton's equation (4.1.3) with Hamiltonian (4.1.4) in t_s up to a trajectory τ.
- A partial momentum refreshment

$$\pi'(t'_s, x) = c_1 \pi(\tau, x) + c_2\, \eta(\tau, x), \tag{4.4.1}$$

  with η a noise with zero mean and variance (4.3.4) and the coefficients $c_1 = e^{-\gamma \delta t}$, $c_2 = (1 - c_1^2)^{1/2}$ that depend on the step size δt and a parameter $\gamma > 0$ which control momentum mixing.

The GHMD has tree free parameters: the trajectory length τ, the momentum mixing coefficient c_1, and the integration step size δt.

If the mixing coefficient c_1 is set to $c_1 = 1$ with an arbitrary value of τ the algorithm reduces to the HMD. The Kramers algorithm is obtained for a trajectory of a single step $\tau = \delta t$ and any value of c_1.

We presented in §4.1.1 an argument that the stationary distribution in HMD differs from the desired one up to order δt^n where n is the order of the integrator. Recently in [46] the range of step size where the algorithm convergence is guaranteed was established in KSPT. In this section we reformulate the proof in [46] and we extend it to the NSPT based on GHMD.

4.4.1 Molecular-dynamics evolution in free theory

The lowest order of a NSPT simulation is a free theory. The dynamics is described by the Hamiltonian (4.1.37) which is a sum of harmonic oscillators with frequencies

ω_p^2: with periodic boundary conditions the ω_p^2 take the form given in (4.1.38). In this case the MD equations are linear and one step δt of their numerical integration amounts to a linear transformation

$$\begin{pmatrix} \phi^{t_0+\delta t}(p) \\ \pi^{t_0+\delta t}(p) \end{pmatrix} = U \begin{pmatrix} \phi^{t_0}(p) \\ \pi^{t_0}(p) \end{pmatrix} \tag{4.4.2}$$

For rest of this section we shall consider only a single oscillator $\omega_p = \omega$, as everything will be block diagonal in p. The reversibility of the integrator requires that $U(-\delta t) = F \circ U(\delta t) \circ F = U(\delta t)^{-1}$ where $F : \pi \to -\pi$ is a momentum flip

$$\begin{pmatrix} U_{1,1}(\delta t) & -U_{1,2}(\delta t) \\ -U_{2,1}(\delta t) & U_{2,2}(\delta t) \end{pmatrix} = \begin{pmatrix} U_{1,1}(-\delta t) & U_{1,2}(-\delta t) \\ U_{2,1}(-\delta t) & U_{2,2}(-\delta t) \end{pmatrix} = \begin{pmatrix} U_{2,2}(\delta t) & -U_{1,2}(\delta t) \\ -U_{2,1}(\delta t) & U_{1,1}(\delta t) \end{pmatrix}. \tag{4.4.3}$$

Area preservation requires $\det U(\delta t) = 1$. Hence, dropping the dependence on δt the reversibility and area preservation implies the following relations

$$U_{1,1} = U_{2,2} \tag{4.4.4}$$

$$U_{1,1}U_{2,2} - U_{1,2}U_{2,1} = 1. \tag{4.4.5}$$

The element $U_{i,j}$ are real polynomials in δt and the frequencies ω^2 (4.1.38). In the case of the leapfrog integrator $U(\delta t)$ take the form in the (4.1.40). It follows from (4.4.4) and (4.4.5) that the Hamiltonian function

$$H(\phi, \pi) = \sum_p \left(\frac{1}{2}\pi(-p)\pi(p) + \frac{1}{2}\hat{\omega}^2\phi(-p)\phi(p) \right) \tag{4.4.6}$$

with

$$\hat{\omega}^2 = \frac{U_{2,1}}{U_{1,2}} = \omega^2 + O(\delta t) \tag{4.4.7}$$

is exactly conserved by one step of the integrator. The eigenvalue of the matrix $U(\delta t)$ are

$$\mu_\pm = U_{1,1} \pm \sqrt{U_{1,1}^2 - 1}. \tag{4.4.8}$$

Depending on the sign of the discriminant in the (4.4.8) there are three cases

1. if the discriminant is negative $U_{1,1}^2 - 1 < 0$ then the eigenvalues are complex conjugates $\mu_+ = \mu$ and $\mu_- = \mu^*$. The condition $\det U(\delta\tau) = 1$ implies $|\mu| = 1$ so the eigenvalues are phases $\mu_\pm = e^{\pm i\theta}$;

2. if the discriminant is zero $U_{1,1}^2 - 1 = 0$ then the eigenvalues are real and degenerate. The condition $\det U(\delta\tau) = 1$ implies $\mu = \pm 1$, so $U(\delta t)$ will be the identity matrix;

3. if the discriminant is positive $U_{1,1}^2 - 1 > 0$ then the eigenvalues are real and $\mu_+ > \mu_-$. The condition $\det U(\delta\tau) = 1$ implies $\mu_+ > 1$ and $\mu_- < 1$.

The time evolution operator $U(\tau) = U(\delta\tau)^{\tau/\delta t}$ for a trajectory.

$$U(\tau) = \begin{pmatrix} U_{1,1}(\tau) & U_{1,2}(\tau) \\ U_{2,1}(\tau) & U_{2,2}(\tau) \end{pmatrix} \tag{4.4.9}$$

will be reversible and area preserving by construction, so it will satisfy the relations

$$U_{1,1}(\tau) = U_{2,2}(\tau) \tag{4.4.10}$$

$$U_{1,1}(\tau)U_{2,2}(\tau) - U_{1,2}(\tau)U_{2,1}(\tau) = 1. \tag{4.4.11}$$

Its eigenvalues $\mu_{\tau\pm}$ are related to the eigenvalue of $\mu_\pm$ as $\mu_{\tau\pm} = \mu_\pm^{\tau/\delta t}$, thus the eigenvalues of $U(\delta\tau)$ belong the same cases as the eigenvalues of $U(\delta t)$, i.e. if $\mu_\pm < 1$ then $\mu_{\tau\pm} < 1$, if $\mu_\pm = 0$ then $\mu_{\tau\pm} = 0$ and if $\mu_\pm > 1$ then $\mu_{\tau\pm} > 1$. The only exception is when the eigenvalues of $U(\delta t)$ are phases $\mu_\pm = e^{\pm i\theta}$ with $\theta\tau/\delta t = \pi N$, where $N = 1, 2, 3, ...$ In this case the eigenvalues of $U(\tau)$ are degenerate and equal to $\mu_\tau = \pm 1$. The time evolution over a trajectory $U(\tau)$ will conserve the Hamiltonian (4.4.6) because it is conserved by a single step $U(\delta t)$, moreover we have that

$$\hat{\omega}^2 = \frac{U_{2,1}}{U_{1,2}} = \frac{U_{2,1}(\tau)}{U_{1,2}(\tau)}. \tag{4.4.12}$$

The latter can be proved by induction. It is true for $\tau = 2\delta t$

$$U(2\delta t) = \begin{pmatrix} U_{1,1}(2\delta t) & U_{1,2}(2\delta t) \\ U_{2,1}(2\delta t) & U_{2,2}(2\delta t) \end{pmatrix} = \begin{pmatrix} U_{1,1} & U_{1,2} \\ U_{2,1} & U_{1,1} \end{pmatrix}^2 \implies \frac{U_{2,1}(2\delta t)}{U_{1,2}(2\delta t)} = \frac{2U_{1,1}U_{2,1}}{2U_{1,1}U_{1,2}}. \tag{4.4.13}$$

We assume the induction hypothesis

$$\frac{U_{2,1}}{U_{1,2}} = \frac{U_{2,1}(\tau)}{U_{1,2}(\tau)} \tag{4.4.14}$$

from this integrating the MD up to $\tau + \delta t$ we get

$$U(\tau+\delta t) = U(\delta t)U(\tau) = \begin{pmatrix} U_{1,1}U_{1,1}(\tau) + U_{1,2}U_{2,1}(\tau) & U_{1,1}U_{1,2}(\tau) + U_{1,2}U_{1,1}(\tau) \\ U_{2,1}U_{1,1}(\tau) + U_{1,1}U_{2,1}(\tau) & U_{2,1}U_{1,2}(\tau) + U_{1,1}U_{1,1}(\tau) \end{pmatrix}, \tag{4.4.15}$$

the ratio of the off-diagonal elements is

$$\frac{U_{2,1}U_{1,1}(\tau) + U_{1,1}U_{2,1}(\tau)}{U_{1,1}U_{1,2}(\tau) + U_{1,2}U_{1,1}(\tau)} = \frac{U_{2,1}\frac{U_{1,1}(\tau)}{U_{1,2}(\tau)} + U_{1,1}\frac{U_{2,1}(\tau)}{U_{1,2}(\tau)}}{U_{1,1} + U_{1,2}\frac{U_{1,1}(\tau)}{U_{1,2}(\tau)}} \tag{4.4.16}$$

using the induction hypothesis (4.4.14) the above equation becomes

$$\frac{U_{2,1}\frac{U_{1,1}(\tau)}{U_{1,2}(\tau)} + U_{1,1}\frac{U_{2,1}}{U_{1,2}}}{U_{1,1} + U_{1,2}\frac{U_{1,1}(\tau)}{U_{1,2}(\tau)}} = \frac{U_{2,1}}{U_{1,2}}\frac{\left(\frac{U_{1,1}(\tau)}{U_{1,2}(\tau)} + \frac{U_{1,1}}{U_{1,2}}\right)}{\left(\frac{U_{1,1}}{U_{1,2}} + \frac{U_{1,1}(\tau)}{U_{1,2}(\tau)}\right)} = \frac{U_{2,1}}{U_{1,2}} \tag{4.4.17}$$

which proves (4.4.12).
Note that if we are in the case $\theta\tau/\delta t = \pi N$ with $N = 1, 2, 3, ...$ the matrix $U(\tau)$ will be the identity matrix and $U_{1,2}(\tau)$ will be zero, the ratio $U_{2,1}(\tau)/U_{1,2}(\tau)$ can not be defined but the Hamiltonian (4.4.6) will be conserved.

In the following we will assume that the discriminant of the characteristic equation of $U(\delta t)$ (4.4.8) is negative

$$U_{1,1}^2 < 1 \tag{4.4.18}$$

and

$$\theta\tau/\delta t \neq \pi N. \tag{4.4.19}$$

Further we will assume

$$U_{1,2} \neq 0. \tag{4.4.20}$$

4.4.2 Convergence of the leading order

Starting from an initial configuration ϕ^0, π^0, after n updates of GHMC we obtained

$$\begin{pmatrix} \phi^{n\delta t} \\ \pi^{n\delta t} \end{pmatrix} = \tilde{U}(\tau)^n \begin{pmatrix} \phi^0 \\ \pi^0 \end{pmatrix} + c_2 \sum_{u=0}^{n-1} \tilde{U}(\tau)^{n-u-1} U(\tau) \begin{pmatrix} 0 \\ \eta^u \end{pmatrix} \tag{4.4.21}$$

where

$$\tilde{U}(\tau) = U(\tau) \begin{pmatrix} c_1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{4.4.22}$$

Using the (4.4.4), (4.4.5) the eigenvalues of $\tilde{U}(\tau)$ can be written as

$$\lambda_\pm = \frac{1}{2} \left((1+c_1) U_{1,1}(\tau) \pm \sqrt{(1+c_1)^2 U_{1,1}(\tau) - 4c_1} \right) \tag{4.4.23}$$

In particular, $|\lambda_\pm| < 1$ thus $\tilde{U}(\tau)$ is a contraction mapping and the first term of the right hand side of (4.4.21) falls with increasing simulation time $t_s = n\tau$. Since all the random noise vectors η are gaussian distributed, the momenta π and the fields ϕ will also be gaussian distributed with mean zero. The variance of the distribution of ϕ is equal to the two-point function that in the large time limit is given by

$$\left\langle \phi^{n\,\delta t} \phi^{m\,\delta t} \right\rangle \underset{n \geq m}{=} \tilde{U}(\tau)^{n-m} K_{1,1}, \tag{4.4.24}$$

with

$$K = c_2^2 \sum_{u=0}^{\infty} \tilde{U}(\tau)^u U(\tau) P_+ \left(\tilde{U}(\tau)^u U(\tau) \right)^T, \qquad P_+ = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{4.4.25}$$

The matrix K satisfies the equation

$$\tilde{U}(\tau) K \tilde{U}(\tau)^T = K - c_2^2 U(\tau) P_+ U(\tau)^T, \tag{4.4.26}$$

which has the unique solution

$$K = \begin{pmatrix} 1 & 0 \\ 0 & \hat{\omega}^{-2} \end{pmatrix}. \tag{4.4.27}$$

Thus the distribution of ϕ simulated by the GHMC algorithm is

$$P[\pi,\phi] \propto e^{-\frac{1}{2}\sum_p \left(\pi(-p)\pi(p)+\hat{\omega}^2\phi(-p)\phi(p)\right)} \tag{4.4.28}$$

provided that the condition (4.4.18), (4.4.19) and (4.4.20) are satisfied.

4.4.3 Convergence beyond the leading order

The structure of the action $S[\phi]$ implies that the force in the MD integration is of the form

$$F_r(\phi_0,\cdots,\phi_r) = \Delta\phi_r + F'_r(\phi_0,\cdots,\phi_{r-1}) \tag{4.4.29}$$

in a theory with only one coupling. For ϕ^4 theory expanding g_0 and δm in x space we have

$$F'_{r,s}(\phi_0,\cdots,\phi_{r-1}) \;=\; \phi_{r,s-1} \;+\; \frac{1}{3!}\sum_{\substack{i_1,\,i_2,\,i_3\\ j_1,\,j_2,\,j_3}} \phi_{i_1,j_1}\phi_{i_2,j_2}\phi_{i_1,j_3}\delta_{i_1+i_2+i_3,r-1}\delta_{j_1+j_2+j_3,s}.$$

If the history of the field $\phi_0,\cdots,\phi_{r-1}$ is known the iteration of n steps of GHMC at order $r>0$ correspond to solving the inhomogeneous linear recursion

$$\begin{pmatrix}\phi_r^{n\delta t}\\ \pi_r^{n\delta t}\end{pmatrix} = \tilde{U}(\tau)^n \begin{pmatrix}\phi_r^0\\ \pi_r^0\end{pmatrix} + c_2\sum_{u=0}^{n-1}\tilde{U}(\tau)^{n-u-1}V_r(\phi_0,\cdots,\phi_{r-1}) \tag{4.4.30}$$

where the vertices $V_r(\phi_0,\cdots,\phi_{r-1})$ are polynomials in their arguments: their exact form depends on the integrator scheme and the forces. Using the fact that $\tilde{U}(\tau)$ is a contraction mapping, the convergence of the high order autocorrelation function can be showed recursively starting from the lowest order. Any high order autocorrelation function can be related with the (4.4.30) to lower order autocorrelation function up to the exponentially decaying contribution

$$\left\langle \sum_{i_1,\cdots,i_k}\phi^n_{i_1}\cdots\phi^n_{i_k}\delta_{i_1+\cdots+i_k,r}\right\rangle = \sum_{m<n} c_m \left\langle \sum_{i_1,\cdots,i_k}\phi^m_{i_1}\cdots\phi^m_{i_k}\delta_{i_1+\cdots+i_n,s}\right\rangle \qquad s<r \tag{4.4.31}$$

with c_m some coefficients. In the large time limit the autocorrelation function do not depend on the initial condition and are stationary and so invariant under time translation

$$\left\langle \sum_{i_1,\cdots,i_k} \phi^m_{i_1} \cdots \phi^m_{i_k} \delta_{i_1+\cdots+i_n,s} \right\rangle = \left\langle \sum_{i_1,\cdots,i_k} \phi^t_{i_1} \cdots \phi^t_{i_k} \delta_{i_1+\cdots+i_n,s} \right\rangle \qquad \forall m,t. \quad (4.4.32)$$

4.4.4 Randomised trajectory length

So far we excluded the case when the integrator of the molecular dynamics equations 4.4.9 have eigenvalues equal to ± 1. In this case the field configuration remain unchanged (or just change sign) during the Monte Carlo history and so the algorithm will not converge.
The condition (4.4.18) provides that a single step of the integrator used in the MD $U(\delta t)$ equations do not have an eigenvalue equal to ± 1, however if the condition (4.4.19) is not fulfilled the integration over a trajectory $U(\tau)$ will have an eigenvalue equal to ± 1. The condition (4.4.19) can be neglected if the trajectory length is randomized [43]. The resulting MD integrator averaging over the distribution of trajectory lengths will be

$$\langle U(\tau) \rangle_\tau = \int d\tau P(\tau) U(\tau) \quad (4.4.33)$$

where $P(\tau)$ is the probability distribution of τ. Since the $U(\tau)$ commute with each other for all values of τ they can be diagonalized simultaneously, so the eigenvalue $\langle \mu_{\tau\pm} \rangle_\tau$ of $\langle U(\tau) \rangle_\tau$ are

$$\langle \mu_{\tau\pm} \rangle_\tau = \int d\tau P(\tau) \mu_{\tau\pm} \quad (4.4.34)$$

where $\mu_{\tau\pm}$ are the eigenvalue of $U(\tau)$. Since $|\mu_{\tau\pm}| = 1$ the real part of $\mu_{\tau\pm}$ have to be $|\mathrm{Re}\, \mu_{\tau\pm}| \leq 1$, so

$$|\mathrm{Re}\ \langle \mu_{\tau\pm} \rangle_\tau\ | \leq \int d\tau P(\tau) |\mathrm{Re}\, \mu_{\tau\pm}|, \quad (4.4.35)$$

if there exists a set of positive measure I where $|\mathrm{Re}\,\mu_{\tau\pm}| < 1$ and $P(\tau) > 0$ we have

$$\begin{aligned} |\mathrm{Re}\ \langle\mu_{\tau\pm}\rangle_\tau\,| &\leq \int_{\mathbb{R}-I} d\tau P(\tau)|\mathrm{Re}\,\mu_{\tau\pm}| + \int_I d\tau P(\tau)|\mathrm{Re}\,\mu_{\tau\pm}| \\ &< \int_{\mathbb{R}-I} d\tau P(\tau) + \int_I d\tau P(\tau) = 1 \end{aligned} \tag{4.4.36}$$

where the strict inequality comes from the integration over I. (4.4.36) implies that $\langle\mu_{\tau\pm}\rangle_\tau \neq 1$ so the discussion in the previous sections applies here without any changes, thus the convergence of the algorithm to a unique stationary state is guaranteed to all orders in the coupling.
In the case that τ is chosen from a discrete distribution probability we will have

$$\langle\mu_{\tau\pm}\rangle_\tau = \frac{\sum_\tau w_\tau \mu_{\tau\pm}}{\sum_\tau w_\tau} \tag{4.4.37}$$

with $w_\tau \geq 0$. If there is an $w_\tau > 0$ corresponding to a $\mu_{\tau\pm} \neq 0$ similar steps as in the 4.4.36 lead to $\langle\mu_{\tau\pm}\rangle_\tau \neq 1$ and thus to the convergence of the algorithm.

4.4.5 Numerical Tests

In the leapfrog (4.1.20) integration scheme the matrix $U(\delta t)$ is

$$U_{1,1} = U_{2,2} = 1 - \frac{1}{2}\omega_p^2\delta t^2\,, \quad U_{1,2} = \delta t\,, \quad U_{2,1} = -\omega_p^2\delta t(1 - \frac{1}{4}\omega_p^2\delta t^2). \tag{4.4.38}$$

The condition that the discriminant of the (4.4.8) is negative implies

$$\delta t^2 < \frac{4}{\omega_p^2}. \tag{4.4.39}$$

In fig. 4.1 the norm of the lowest order field is plotted during the Monte Carlo history in a 4^4 box with $m = 1$ for HSPT. In this case the highest frequency ω_p^2 (4.1.38) is equal to $\omega_{p,max}^2 = 17$, thus the GHMD algorithm converges if $\delta t^2 < 0.485$ for all values of γ and $\langle\tau\rangle$. For $\delta t = 0.49$ (left panel) the norm of the lowest order field grows very rapidly during the Monte Carlo evolution for different values of

$\langle\tau\rangle$, this is a signature of the instability of the algorithm. While for $\delta t = 0.48$ (right panel) the norm of the field appears to be stable for all the three values of the average trajectory length.

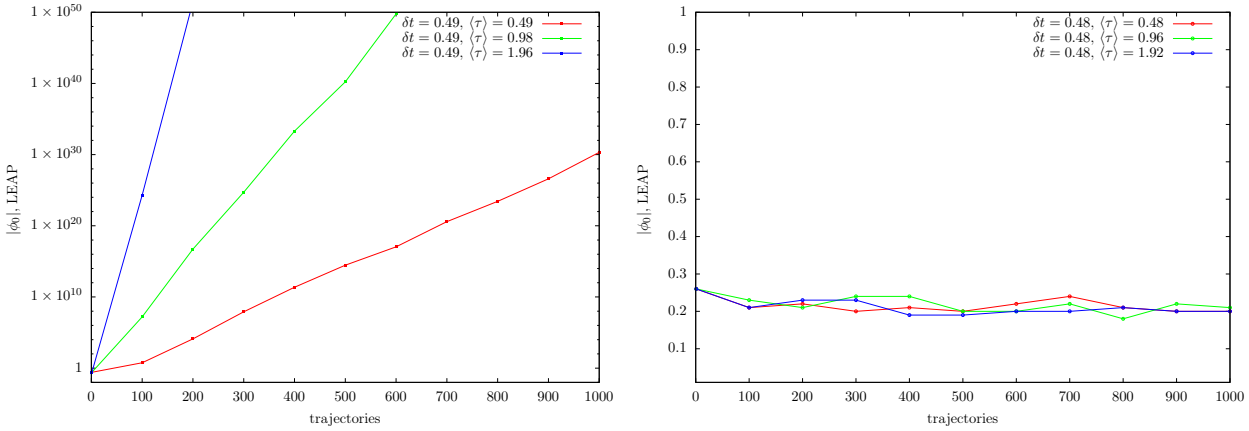


Figure 4.1 *Norm of the lowest-order field during the Monte Carlo history of an HSPT simulation in a 4^4 box with $m = 1$ with the leapfrog integration scheme (LEAP). The algorithm is expected to be unstable for $\delta t = 0.485$*

Chapter 5

An implementation of ISPT in φ^4 theory

Here we present a new technique ISPT introduced in [16] [1], this form of stochastic perturbation theory is unrelated to stochastic quantization and is instead based on a diagrammatic expansion of the trivializing map. In the case of ISPT the stochastic field is given by a series of trees that can be computed through a recursive procedure. Here we will describe the essential features of this approach so as to emphasize the most prominent differences with standard NSPT techniques.

5.0.1 Definitions

ISPT is based on the concept of trivializing maps. In the most general case these transform a set of Gaussian-distributed random fields $\eta_i(x)$, for $i = 0, 1, 2, \ldots$, into a stochastic field $\phi(x)$ such that

$$\big\langle \phi(x_1) \cdots \phi(x_n) \big\rangle_\eta = \big\langle \varphi(x_1) \cdots \varphi(x_n) \big\rangle \tag{5.0.1}$$

[1] Additional useful material is provided by the author of [16] in the documentation for the publicly available package [48].

order-by-order in the couplings of the theory. Here the expectation value on the right hand side is defined by (2.0.4), whereas on left hand side it is given in terms of averages over the Gaussian random fields

$$\big\langle \eta_i(x) \big\rangle_\eta = 0, \qquad \big\langle \eta_i(x)\eta_j(y) \big\rangle_\eta = \delta_{ij}\delta_{xy}. \tag{5.0.2}$$

In perturbation theory the stochastic field ϕ can be represented as a power series in the couplings of the theory as in (3.5.3). In particular, in the regularized theory we can consider an expansion in terms of the bare coupling g_0,

$$\phi(x) = \sum_{k=0}^{N} \phi_k(x) g_0^k + O(g_0^{N+1}). \tag{5.0.3}$$

Given this the corresponding expansion in terms of a renormalized coupling is easily obtained using relation (2.0.3) (q.v., Appendix A.1). On the other hand, the determination of the coefficients ϕ_k in terms of the renormalized instead of the bare mass requires explicit computation of the mass counterterm contributions. As before, in the numerical implementation of the method it is convenient to store the field as a two-dimensional array $\phi_{k,\ell}$ with the indices corresponding to the powers of g_0 and δm^2 (3.5.4)

$$\phi(x) = \sum_{k,\ell=0}^{N} \phi_{k,\ell}(x) g_0^k (\delta m^2)^\ell + O(g_0^{N+1}). \tag{5.0.4}$$

Once the expansion (2.0.2) is known it is trivial to pass from the representation (3.5.4) to (3.5.3). Using the representation (3.5.4) the expansion (2.0.2) can be determined and thus the results obtained in terms of the renormalized mass. This is discussed in detail in Appendix A.2.

We find at the lowest-order in the coupling

$$\phi_{0,0}(x) = \sum_{y\in\Omega} H(x,y)\eta_0(y), \tag{5.0.5}$$

where H is the Greens function for the operator $\sqrt{-\partial^2 + m^2}$,

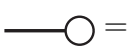
$$H(x,y) = \frac{1}{L^4} \sum_{p \in \tilde{\Omega}} e^{ip(x-y)} \sqrt{\tilde{G}(p)}, \qquad \text{where} \qquad \tilde{G}(p) = \frac{1}{\hat{p}^2 + m^2}. \tag{5.0.6}$$

It is easy to show that this field satisfies (5.0.1) at lowest order in the coupling.

Beyond the leading order there is more freedom to define the trivializing field. Following [16] we write this as a linear combination of the values $v(x, \mathcal{R}_i)$ of the rooted tree diagrams $\mathcal{R}_i$ with coefficients $c(\mathcal{R}_i)$,

$$\phi_{k,l}(x) = \sum_{i \in \mathcal{S}_{k,l}} c(\mathcal{R}_i) v(x, \mathcal{R}_i), \tag{5.0.7}$$

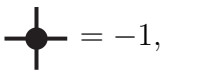
where $\mathcal{S}_{k,l}$ is the set of all diagrams of order g_0^k and $(\delta m^2)^l$. Graphical representations of the rooted tree-diagrams contributing to $O(g_0)$ $(k + \ell = 1)$ and $O(g_0^2)$ $(k + \ell = 2)$ are given in Figures 5.1 and 5.2 respectively; the corresponding coefficients $c(\mathcal{R}_i)$ are also shown. In this representation the leaves of the trees are given by



$$= \chi_i(x) = \sum_{y \in \Omega} H(x,y) \eta_i(y), \tag{5.0.8}$$

where the index i is the number adjacent to the open circle in the graph; if no such number is displayed it is implicit that $i = 0$. The leaves are thus given by the lowest order solution (5.0.5) with the appropriate choice of random field η_i.

Black circles and crosses represent the vertices of the theory: they are the usual φ^4 vertex and mass counterterm insertions,



$$= -1, \qquad = 1. \tag{5.0.9}$$

These are associated with implicit factors of $-g$ and δm^2 respectively. Black lines connecting two vertices correspond to the scalar propagator,

$$= G(x,y) = \frac{1}{L^4} \sum_{p \in \tilde{\Omega}} e^{ip(x-y)} \tilde{G}(p), \tag{5.0.10}$$

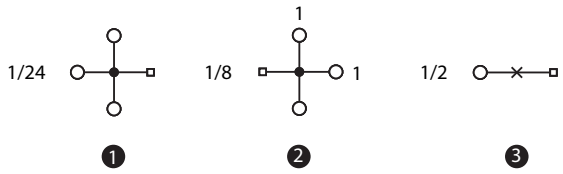


Figure 5.1 *Rooted tree-diagrams contributing at $O(g_0)$; note that $\delta m^2 = O(g_0)$.*

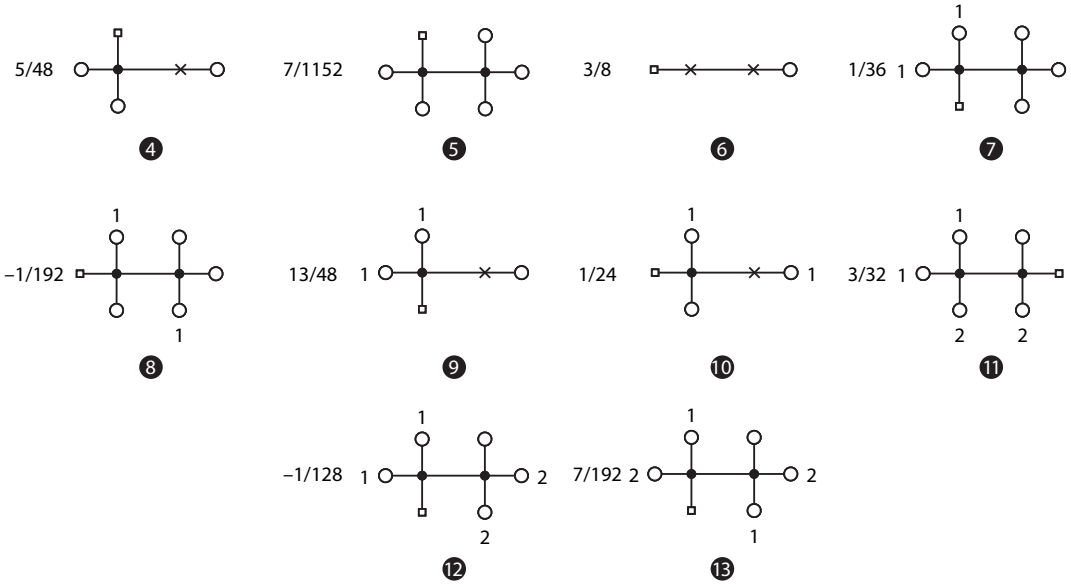


Figure 5.2 *Rooted tree-diagrams contributing at $O(g_0^2)$; note that $\delta m^2 = O(g_0)$.*

where x and y are the positions of the two vertices connected by the given propagator. In particular, at each vertex the fields attached are multiplied together and the propagator is applied to the resulting product of fields.

The root of the diagram is given by

$$= G(x, y), \tag{5.0.11}$$

where x is the space-time index of the corresponding rooted tree $\mathcal{R}_i$.

To give some examples, given some $\eta_0(x)$ and $\eta_1(x)$ fields, the diagram labelled 2 in Figure 5.1 evaluates to

$$v(x, \mathcal{R}_2) = (-1) \sum_{y \in \Omega} G(x, y) \chi_0(y) \chi_1(y)^2. \tag{5.0.12}$$

This contributes to $\phi_{1,0}(x)$ with a coefficient $c(\mathcal{R}_2) = 1/8$. Diagram 4 in Figure 5.2, stands for

$$v(x, \mathcal{R}_4) = (-1) \sum_{y \in \Omega} G(x, y) \chi_0(y)^2 \sum_{z \in \Omega} G(y, z) \chi_0(z), \tag{5.0.13}$$

and contributes to $\phi_{1,1}(x)$ with $c(\mathcal{R}_4) = 5/48$.

g_0^k	n	$c(\mathcal{R}) \neq 0$
1	3	3
2	10	10
3	44	43
4	241	231
5	1,506	1,420
6	10,778	10,015
Total	12,582	11,722

Table 5.1 *Number of rooted tree diagrams appearing at a given order in the coupling g_0. The column labeled by $c(\mathcal{R}) \neq 0$ gives the number of such diagrams whose coefficient $c(\mathcal{R}_i)$ is non-vanishing.*

Given these examples it is clear that the evaluation of the trivializing map for a given set of random fields η_i is in principle straightforward. Beyond the lowest

perturbative orders the number of diagrams (as well as their complexity) increases rapidly as indicated in Table 5.1, so the computation must be automated.

For this work we wrote a program that evaluates the trivializing field $\phi(x)$ up to an arbitrary order N in the couplings for a given set of η_i fields. For the structure of the relevant diagrams and the determination of their coefficients we used the software package provided by Martin Lüscher [48]. The diagrams are given as `C structs` of abstract elements, so our program visits each vertex in a diagram using depth-first recursion starting from the root, and evaluates the corresponding numerical expressions. The diagrams are collected according to their order in the couplings and the $\phi_{k,l}(x)$ fields are thus constructed. This allows the series (3.5.4) to be obtained for some set of η_i fields. Once this is done, as in LSPT correlation functions of the trivializing field can be computed order by order in the couplings with the automated operations of the form (3.5.6) and evaluated stochastically by averaging over different samples of the Gaussian random fields η_i.

We should mention some additional technical details. First, in the diagrammatic computation the scalar propagators are applied in momentum space, while the product of fields at vertices are performed in position space. This is implemented using the efficient numerical evaluation of the discrete Fourier transformation provided by the FFTW package [49]. As a result the cost of the computation of the diagrams scales proportionally to the system size $V = L^4$ up to logarithms. Second, as already noted in [16], the computation of the rooted tree diagrams could be organized in such a way that identical sub-trees in different graphs are cached. How to do this efficiently is a non-trivial issue even for φ^4 theory, and we did not investigate it further because, as we shall see below, ISPT suffers for some severe limitations once high-order computations are considered. Its utility might thus be limited to relatively low-order computations where recomputation of subgraphs is not a significant issue.

The advantages of ISPT is that its results are exact up to statistical uncertainties and that there are no autocorrelations as the coefficients $\phi_{k,\ell}$ are generated "instantaneously" from independent Gaussian random fields η_i.

5.0.2 A test of the method

We tested our ISPT implementation by comparing some results with those obtained using conventional perturbative lattice calculations (LPT). We computed the renormalized coupling (2.1.9) and compared it with its two-loop determination from [17], which we evaluated for the parameters of interest (see below). We considered both the case where the perturbative expansion is given in terms of the renormalized mass (2.1.1), and the case where it is given in terms of the bare mass m_0.[2] The comparison was done on a tiny lattice with $L = 4$, where high statistics could be gathered, and the value of the mass was chosen such that $z = 4$. The results of the tests are reported in Table 5.2; for completeness we also give the results for δm^2 in the table.

	Mass	$c_2 \times 10^2$	$c_3 \times 10^3$	$m_1^2 \times 10^2$	$m_2^2 \times 10^4$
LPT	m	-3.330	1.583	-6.4221	3.6702
ISPT	m	$-3.332(6)$	1.582(4)	$-6.4220(1)$	3.6704(6)
LPT	m_0	-3.33	2.965		
ISPT	m_0	$-3.33(1)$	2.964(5)		

Table 5.2 *Results for the series (2.0.3) and (2.0.2) as obtained from ISPT and conventional LPT for $L = 4$ and $z = \text{Mass} \times L$ using 10^8 field configurations. The perturbative expansion for the coupling (2.0.3) is obtained both in terms of the renormalized mass m of (2.1.1) as well as the bare mass m_0.*

As can be seen from the table there is good agreement between the ISPT and the LPT determinations, thus confirming the correctness of our implementation. In the case where the mass renormalization is considered one needs to take into account the effect of statistical errors in the mass renormalization procedure discussed in Appendix A.2: we did this using the jackknife method.

[2] In ISPT the latter is simply obtained by setting $\delta m^2 = 0$ in the expansion (3.5.8).

5.0.3 One loop example

As an example here we will show how to compute the one loop two-point function with ISPT. The two-point function is defined as

$$\langle\phi(x)\phi(y)\rangle_{\eta_0,\eta_1}\big|_{g_0} = \langle\phi_0(x)\phi_1(y)\rangle_{\eta_0,\eta_1} + \langle\phi_1(x)\phi_0(y)\rangle_{\eta_0,\eta_1}.$$

ϕ_0 is defined in the (5.0.5), while ϕ_1 is defined as the sum of the diagrams in figure 5.1 with their coefficient, which is equal to

$$\phi_1(x) = -\frac{1}{24}\sum_{z\in\Omega} G(x,z)\chi_0(z)^3 - \frac{1}{8}\sum_{z\in\Omega} G(x,z)\chi_0(z)\chi_1(z)^2 + \frac{m_1^2}{2}\sum_{z\in\Omega} G(x,z)\chi_0(z) \tag{5.0.14}$$

where m_1^2 is the coefficient of order g_0 of the mass counterterm δm^2 (2.0.2). Since ϕ_0 does not depend on η_1 the (5.0.3) become

$$\langle\phi(x)\phi(y)\rangle_{\eta_0,\eta_1}\big|_{g} = \langle\phi_0(x)\langle\phi_1(y)\rangle_{\eta_1}\rangle_{\eta_0} + \langle\langle\phi_1(x)\rangle_{\eta_1}\phi_0(y)\rangle_{\eta_0}.$$

The average of ϕ_1 respect to the noise η_1 gives

$$\langle\phi_1(x)\rangle_{\eta_1} = -\frac{1}{24}\sum_{z\in\Omega} G(x,z)\chi_0(z)^3 - \frac{1}{8}\sum_{z\in\Omega} G(x,z)\chi_0(z)\Sigma(m) + \frac{m_1^2}{2}\sum_{z\in\Omega} G(x,z)\chi_0(z). \tag{5.0.15}$$

where

$$\Sigma(m) = \langle\chi_1(y)\chi_1(y)\rangle_{\eta_1} = \frac{1}{L^4}\sum_{p\in\tilde{\Omega}} \tilde{G}(p). \tag{5.0.16}$$

The average over η_0 gives

$$\begin{aligned}
&\langle\langle\phi_1(x)\rangle_{\eta_1}\phi_0(y)\rangle_{\eta_0} \\
&= -\frac{3}{24}\sum_{z\in\Omega} G(x,z)\Sigma(m)G(y,z) - \frac{1}{8}\sum_{z\in\Omega} G(x,y)\Sigma(m)G(y,z) + \frac{m_1^2}{2}\sum_{z\in\Omega} G(x,z)G(y,z) \\
&= (-\frac{1}{4} + \frac{m_1^2}{2})\sum_{z\in\Omega} G(x,z)\Sigma(m)G(y,z),
\end{aligned} \tag{5.0.17}$$

which in momentum space become

$$\langle\langle\phi_1(p)\rangle_{\eta_1}\phi_0(q)\rangle_{\eta_0} = \frac{1}{L^8}\sum_{p,q} e^{-ipx-iqy}\langle\langle\phi_1(x)\rangle_{\eta_1}\phi_0(y)\rangle_{\eta_0}$$
$$= (-\frac{1}{4}+\frac{m_1^2}{2})\frac{\delta_{p+q,0}}{L^4}\tilde{G}(p)^2\Sigma(m). \tag{5.0.18}$$

Since the last equation is invariant under interchange of p with q we have that

$$\langle\phi(p)\phi(q)\rangle_{\eta_0,\eta_1}\big|_g = 2\langle\langle\phi_1(p)\rangle_{\eta_1}\phi_0(q)\rangle_{\eta_0} = (-\frac{1}{2}+m_1^2)\frac{\delta_{p+q,0}}{L^4}\tilde{G}(p)^2\Sigma(m). \tag{5.0.19}$$

which is the two point function in φ^4 theory.

After computing the two-point function we want to focus on the role of the diagram 2 of figure 5.1. To do so first we redo the calculation above replacing the coefficients of the diagrams 1 and 2 of figure 5.1 with generic values $c(\mathcal{R}_1) = A$ and $c(\mathcal{R}_2) = B$ obtaining

$$\langle\phi(p)\phi(q)\rangle_{\eta_0,\eta_1}\big|_g = (6A+2B+m_1^2)\frac{\delta_{p+q,0}}{L^4}\tilde{G}(p)^2\Sigma(m). \tag{5.0.20}$$

Imposing that the latter will be equal to the (5.0.19) gives the equation

$$6A+2B = -\frac{1}{2}\,, \tag{5.0.21}$$

which is not enough to determine A and B. The coefficient A and B have to be determined imposing the (5.0.1). The connected and disconnected parts of the correlation functions are related to each other through the moment-cumulant transformation and thus in a way that does not refer to any particular properties of the correlation functions. Equation (5.0.1) is therefore satisfied if and only if the connected parts of the correlation functions are the same. In φ^4 theory at order g_0 the only connected correlation function, apart the two-point function is the four-point function. The computation of the connected four-point function in

ISPT gives

$$\langle\phi(p_1)\phi(p_2)\phi(p_3)\phi(p_4)\rangle_{\eta_0,\eta_1}\big|_g = 4(3!)A\frac{\delta_{p_1+p_2+p_3+p_4,0}}{L^4}\tilde{G}(p_1)\tilde{G}(p_2)\tilde{G}(p_3)\tilde{G}(p_4). \tag{5.0.22}$$

Imposing that the latter will be equal to the connected four-point function of φ^4 theory gives the equation

$$4(3!)A = -1. \tag{5.0.23}$$

The equations (5.0.21) and (5.0.23) have as unique solution $A = -1/24$ and $B = -1/8$. Since we match all the connected correlation functions of the theory at order g_0 all the correlation function of ISPT satisfy the (5.0.1) at order g_0.

Chapter 6

Numerical results

This chapter contain the result of the numerical investigation of the methods described in §3.5—§5. The aim of the numerical investigation is to identify the principal advantages and disadvantages of each technique. In §6.0.1 we compare the perturbative results for some specific quantities obtained with the different algorithms, in order to confirm their correctness and viability. Once these are established, in §6.0.2—§6.0.5 we study the continuum scaling of the errors of these perturbative coefficients as computed by the various methods.

For this numerical investigation we use the ECDF cluster (eddie). The different implementation of NSPT algorithms we wrote are all serial codes. We run the KSPT, ISPT and LSPT simulations on single core Intel Xeon E5-2630 Processor (2.4 GHz) of the ECDF cluster (eddie) where we use approximately 75000 core hours to produce all the result of KSPT, 15000 for ISPT and less than 2000 for LSPT. The HSPT simulations were done using the computer center at DESYZeuthen, where we use 55000 core hours.

Results of our study appeared also in [1] and [2]

6.0.1 Testing the methods

Before comparing the efficiency of the various algorithms it is necessary to confirm that they agree for the perturbative computation of some quantities. In Figure 6.1 the results for $\mathcal{E}(t)$ (2.1.12) at tree-level, $O(g_0)$, and $O(g_0^2)$ are shown from top to bottom respectively. The computations were performed on a tiny $L = 4$ lattice for which very high statistics could be obtained. We collected $\approx 10^7$ independent measurements for ISPT, HSPT, and KSPT, and $\approx 10^6$ measurements for each of 9 values of $\varepsilon \in [0.01, 0.05]$ for LSPT. The values for the mass of the field and the flow time were chosen to correspond to $z = 4$ and $c = 0.2$, respectively (defined in §2.1). For HSPT and KSPT we then chose $\langle \tau \rangle = 1$ and $\gamma = 2$. The perturbative expansion is expressed in terms of the renormalized mass m whereas the perturbative coefficients correspond to the expansion in the bare coupling g_0, i.e.,

$$\mathcal{E}(L, z, c) = \mathcal{E}_0 + \mathcal{E}_1 g_0 + \mathcal{E}_2 g_0^2 + \mathcal{E}_3 g_0^3 + O(g_0^4), \quad \text{where} \quad \mathcal{E}_i \equiv \mathcal{E}_i(L, z, c). \quad (6.0.1)$$

As can be seen from the figure, all the methods agree with each other and with the analytic determination. In the case of LSPT deviations from the expected results are sizable at the largest step-sizes, and agreement is found only after extrapolation to $\varepsilon \to 0$. In particular, the asymptotic $O(\varepsilon^2)$ behaviour expected for the integrator used is clearly visible.

For the case of HSPT and KSPT we do not see any indication of step-size errors as the results show no statistically significant deviation from the analytic determination and ISPT; the points are precise at the 0.1–0.5% level depending on the order. Even though the lattice is quite small, the step-size we chose for both HSPT and KSPT is rather large, namely $\delta t = 0.5$. This step-size satisfies $\delta t^4 \geq 25\,\varepsilon^2$, for all values of ε considered for LSPT: this inequality would give the naïve size of the expected relative step-size errors. However the systematic error in HSPT and KSPT is significantly smaller than LSPT, Consequently, it is feasible to run the algorithm with a step-size such that extrapolations are not required. The magnitude of the systematic error depends on many factors: the lattice size considered, the observable, the parameters of the theory, the values

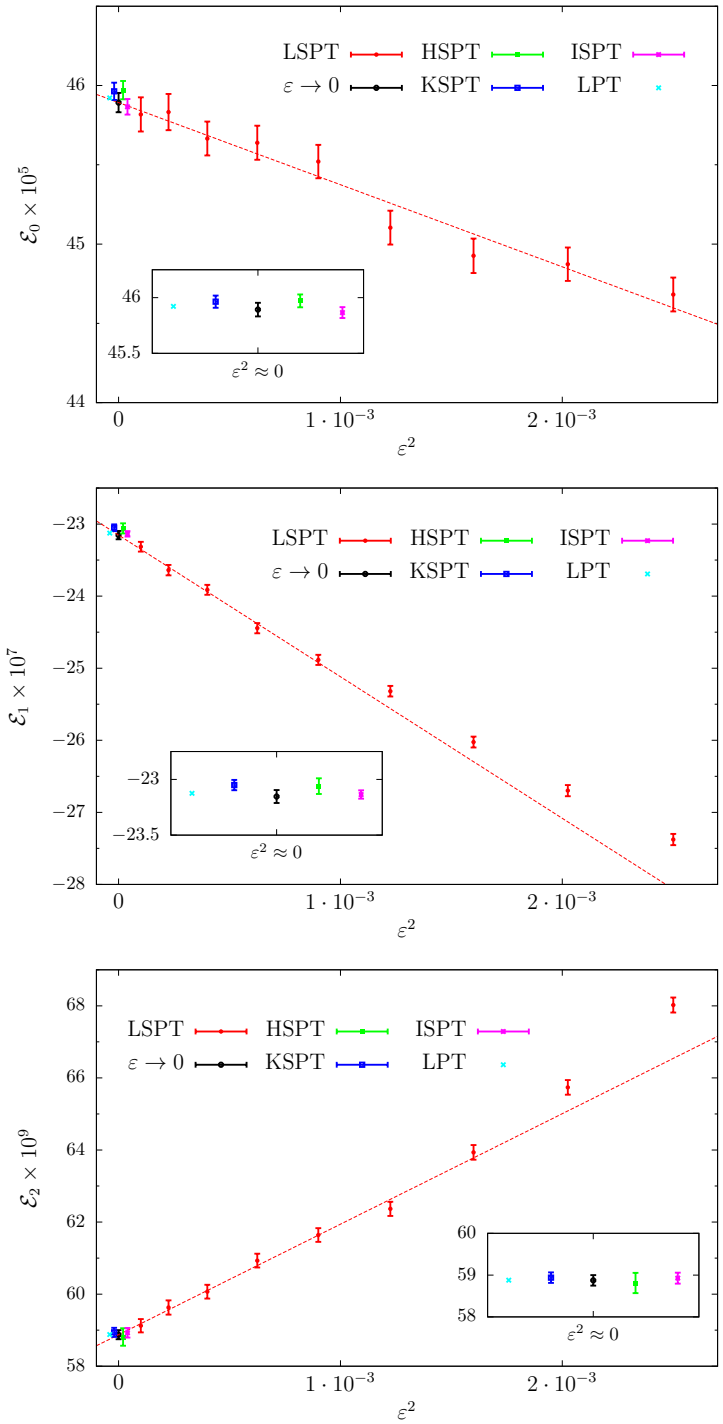


Figure 6.1 *Comparison of different methods in the determination of* $\mathcal{E}_0$, $\mathcal{E}_1$ *and* $\mathcal{E}_2$ *for* $z = 4$, $c = 0.2$, *and* $L = 4$. *The analytic result (LPT) and the result of the extrapolation* $\varepsilon \to 0$ *for LSPT, as well as the ISPT, KSPT, and HSPT results (for which there are no step-size errors or the step-size errors are negligible compared with the statistical errors) are plotted near* $\varepsilon^2 = 0$.

of the step-sizes, and most importantly the integrators used. Nonetheless, as already emphasized, symplectic MD integrators are at a more mature stage of development than Runge Kutta integrators; they can be optimized to reduce the magnitude of the step-size errors as in the case of OMF4 integrator (q.v., [45]), and as illustrated in the example above, this results in a significant reduction of systematic errors relative to the cost of a single integration step. Moreover, as we can afford to run with larger step-sizes for a fixed systematic error and with a fixed number of force computations the cost of obtaining independent configurations is reduced because of the smaller autocorrelations. This needs to be compared with the fact that we measured that the computer time needed for an application of the OMF4 integrator is twice than the RK2 integrator. This is compatible with the expectation that the cost is dominated by the force computation: the OMF4 requires six force computation [45] while the RK2 requires three force computations §3.4.19.

6.0.2 Continuum error scaling: a first look

Having addressed the issue of systematic errors, we now study the continuum scaling of the various NSPT algorithms. We do this by investigating how the (relative) errors of the perturbative coefficients of some given observables scale as the continuum limit of the theory is approached. The precise details of the scaling depend on the observable, but some general features may be inferred.

In Figures 6.2 and 6.3 we show the continuum scaling of the relative errors $\Delta\mathcal{E}_i/\mathcal{E}_i$ for $i = 0, \ldots, 3$ and L in the range $4 \leq L \leq 16$, as computed using ISPT, LSPT, HSPT, and KSPT. Recall that the error $\Delta\mathcal{E}_i$ may be expressed as in (B.2.13)

$$\Delta\mathcal{E}_i = \sqrt{\frac{2A_I(E_i) \times \mathrm{Var}(E_i)}{N}}, \tag{6.0.2}$$

where E_i is

$$\frac{t^2}{L^4} \sum_{x \in \Omega} E(t, x) = E_0 + E_1 g_0 + E_2 g_0^2 + E_3 g_0^3 + O(g_0^4). \tag{6.0.3}$$

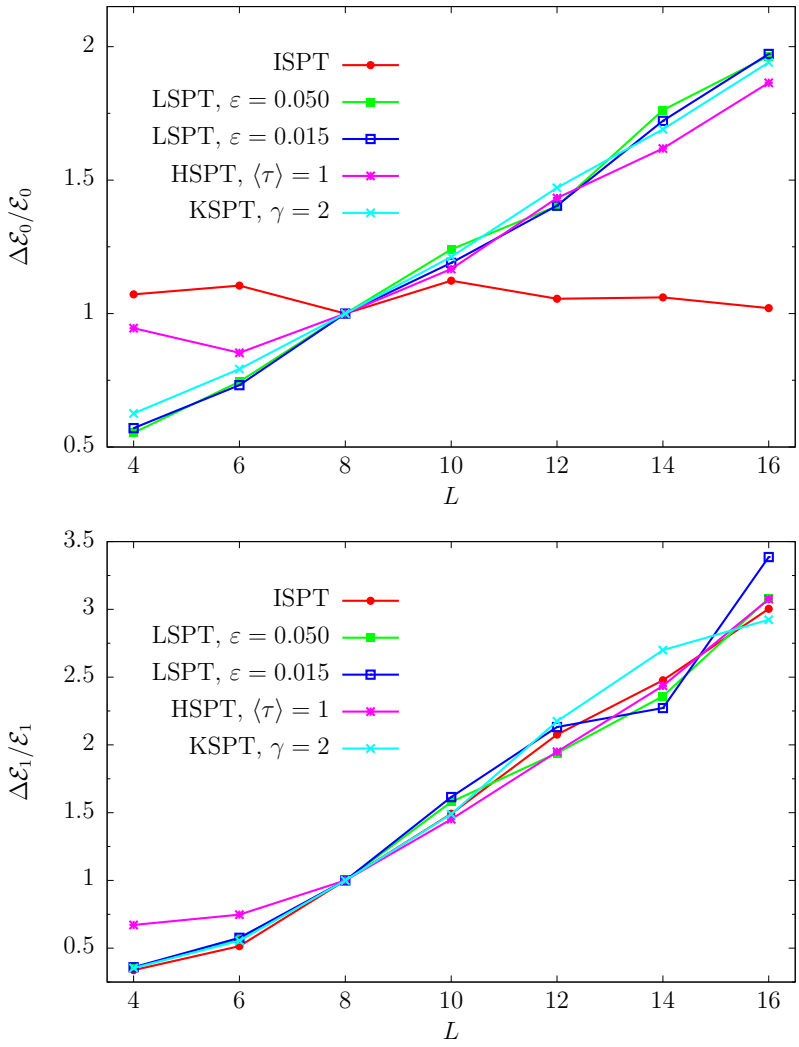


Figure 6.2 *Continuum scaling of the relative errors* $\Delta\mathcal{E}_0/\mathcal{E}_0$ *and* $\Delta\mathcal{E}_1/\mathcal{E}_1$ *as computed with ISPT, LSPT, HSPT, and KSPT. The parameters are* $z = 4$ *and* $c = 0.2$*. The data is normalized at* $L = 8$*.*

We show our results for $z = 4$ and $c = 0.2$, but the same qualitative behaviour is observed in other cases. The number of configurations for each method is specified at $L = 4$ and kept constant as $1/L \to 0$. Specifically, at $L = 4$ we collected between $N = 10^5$ and $N = 10^6$ independent measurements for each of the different methods. At this small lattice size and for the algorithmic parameters considered the different methods have comparable statistical precision for the same number of independent measurements. In the case of LSPT we measured after each step of the Markov chain. For KSPT we set the parameter $\gamma = 2$, and we adjusted the measurement frequency so as to measure at fixed intervals $\Delta t_s = 0.5$ of simulation time independent of the step-size.[1] For HSPT we measured after each trajectory of average length $\langle \tau \rangle = 1$. The results in the figures are normalized to the values of the relative errors at $L = 8$, and hence to a first approximation are independent of N. Since the figures are only intended to be qualitative no estimates for the error on the relative error are provided.

The error computation for the perturbative coefficients was obtained using jackknife in the case of ISPT, whereas for LSPT, HSPT, and KSPT we employed the Γ-method described in Appendix B in order to take into account autocorrelations of the measured quantities. The coefficients $\mathcal{E}_i$ and corresponding errors refer to the expansion in terms of the renormalized mass m and bare coupling g_0 (6.0.1). Power divergences in the inverse lattice spacing are thus excluded in the coefficients $\mathcal{E}_i$, while logarithmic divergences associated with renormalization of the coupling constant are not expected to be relevant for the following discussion. In the case of HSPT and KSPT the step-size was scaled as $\delta t \propto 1/L$ starting from a value of $\delta t = 0.5$ so as to keep the $O(\delta t^4)$ errors in the equilibrium distribution approximately fixed using the OMF4 integrator as the continuum limit is approached (q.v., §4.4), this is probably a very conservative choice, but it was done to avoid potentially large systematic errors that might modify the overall picture. In fact with this choice the step-size errors vanish faster than the leading $O(1/L^2)$ lattice artifacts as the continuum limit is approached. Keeping the systematic errors in the equilibrium distribution fixed for LSPT is significantly more challenging as it requires $\epsilon \propto 1/L^2$ with the RK2 integrator

[1] Since autocorrelations are linear in the step-size δt for γ fixed, from the point of view of autocorrelations this is equivalent to measuring after each step for a fixed step-size of $\delta t = 0.5$.

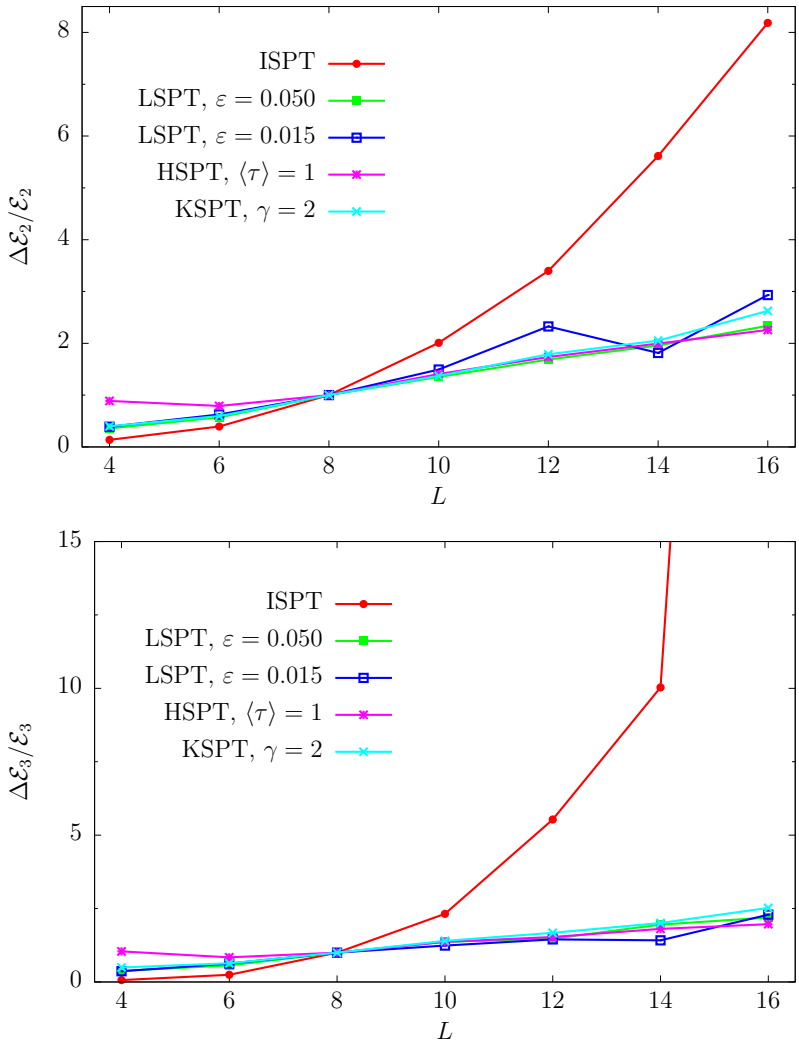


Figure 6.3 *Continuum scaling of the relative errors* $\Delta\mathcal{E}_2/\mathcal{E}_2$ *and* $\Delta\mathcal{E}_3/\mathcal{E}_3$ *as computed with ISPT, LSPT, HSPT, and KSPT. The parameters are* $z = 4$ *and* $c = 0.2$*. The data is normalized at* $L = 8$*. Note that for ISPT* $\Delta\mathcal{E}_3/\mathcal{E}_3 \approx 65$ *for* $L = 16$*.*

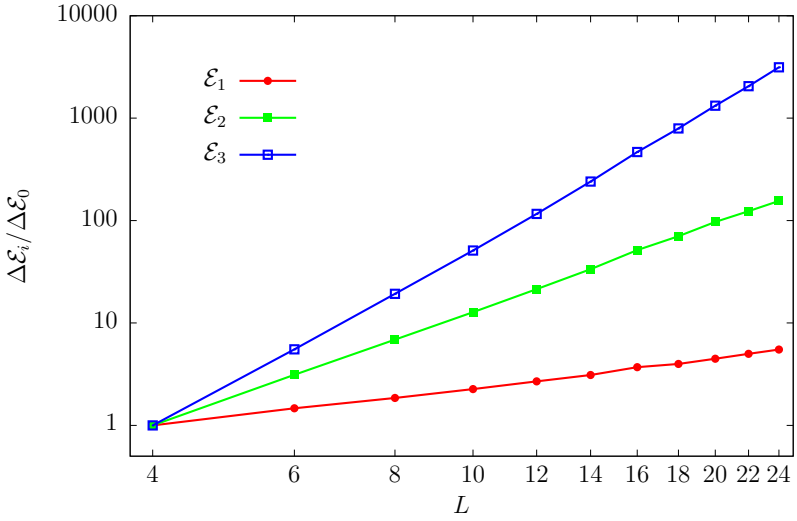


Figure 6.4 *Continuum scaling of the ratios* $\Delta\mathcal{E}_i/\Delta\mathcal{E}_0$ *for* $i = 1, 2, 3$ *as computed with ISPT. The case with* $z = 4$ *and* $c = 0.2$ *is shown. The results are normalized to their values at* $L = 4$*.*

(q.v., §3.4). In this case we thus simply considered two well-separated step-sizes in order to assess the dependence of the results on ϵ.

Starting from the results at tree-level (top panel in Figure 6.2) we see how the relative error of ISPT is constant for a fixed number of field configurations. The results for LSPT, HSPT, and KSPT are rather different: excluding perhaps the smaller lattices there is a linear growth of the relative errors with the lattice size. These results confirm free field theory expectations. The variance $\mathrm{Var}(E_0)$ is finite and constant with L up to discretization effects. In particular it is the same for all NSPT methods up to step-size errors, and independent of the algorithmic parameters. Consequently in ISPT the error $\Delta\mathcal{E}_0$ is essentially constant with L for a given number of field configurations N, because the configuration are uncorrelated. In LSPT, HSPT and KSPT the result are correlated, the integrated autocorrelations A_i grow $\propto L^2$ as the continuum limit is approached and thus the error at tree level will be $\propto L$.

At higher perturbative orders the situation for ISPT changes significantly. At $O(g_0)$ the relative error grows linearly with L, indicating a growth of the variance proportional to L^2 as the continuum limit is approached. For higher

perturbative orders the increase of the variance is even more rapid. This may be better appreciated from Figure 6.4 where results for ISPT alone are given up to $L = 24$ and $O(g_0^3)$. In this plot we show the ratios of $\Delta\mathcal{E}_i$ and $\Delta\mathcal{E}_0$ for $i = 1, 2$ and 3. These ratios are independent on the number of configurations considered, and were estimated using 10^5–10^7 measurements, depending on the lattice size. It is clear that the error, and hence the variance, increases as an increasing power of L as the perturbative order is increased.[2] This is to be compared with the relative errors for LSPT, HSPT, and KSPT, which have the same qualitative behaviour as at tree level, namely the errors increase only linearly with L (Figures 6.2 and 6.3). For these algorithms the behaviour is similar to what happens at tree level: the errors of the higher order coefficients appear to increase due to increasing autocorrelations. The increase of the variance of the perturbative coefficients in LSPT, HSPT, and KSPT, if any, is very mild here.[3] These conclusions will be confirmed by the detailed investigations of the following sections.

We conclude by noticing that the above observations for the higher order results are in agreement with general theoretical expectations. The peculiar behaviour in the variance of perturbative coefficients computed with ISPT was elucidated by Martin Lüscher [50]. In Appendix C we give an example of an observable whose its variance diverges in ISPT and not in LSPT. On the other hand as showed in §3.5.2 and [33] the variances of perturbative coefficients computed using LSPT are at most logarithmically divergent. However, their autocorrelations grow with the square of the correlation length of the system, i.e., $\propto 1/m^2 \propto L^2$. These results cannot be extended to HSPT and KSPT due the non-renormalizability of the HMD equations [26] except for the cases where the algorithm reduces to LSPT, $\tau = \delta t$ for HSPT and $\gamma \to \infty$ for KSPT. These results also apply to KSPT at fixed γ where the Langevin limit is approached together with the continuum limit (q.v., §4.3 and [26]). It is most plausible that they hold in the case where the trajectory length does not scale with the correlation length of the system, i.e., $\langle\tau\rangle$ is independent of L. This follows from the

[2] A similar behaviour was also observed by Martin Lüscher in pure SU(3) Yang-Mills theory [50].

[3] We note that for LSPT a similar observation was made in [12] in the pure SU(3) Yang-Mills theory.

observation that in the continuum limit $L \to \infty$ the HSPT algorithm effectively integrates the perturbatively expanded Langevin equation, as in this case there is no fundamental difference from a single step HSPT algorithm (which is LSPT). This conjecture seems to be confirmed by the numerical experiments discussed below.

6.0.3 Continuum scaling of autocorrelations

As a result of the investigation of the previous subsection we conclude that NSPT methods based on stochastic differential equations have a much better continuum cost scaling than ISPT. It is clear that beyond the first few orders in perturbation theory the scaling of ISPT is such that its performance is much worse than the other algorithms. In this and the following subsection we therefore focus our attention on these other methods. In particular the question we want to address is the following. As is well known, free field analysis of the HMD and Kramers algorithms shows that their continuum cost scaling depends on how their parameters are adjusted [41]. In the context of NSPT these results directly apply to the lowest order determinations. However, it is not obvious what the behaviour of the higher order results is if different parameter scalings are considered; this is because we do not have analytic control on this behaviour except in the Langevin limit of these algorithms. To answer this question we investigate the continuum scaling of the autocorrelations of the perturbative orders E_i as a function of the algorithmic parameters in this subsection. More precisely, we will compare the optimal parameter scaling suggested by the free field theory analysis of [41] with the Langevin scaling. We identify the latter as the case where $\langle\tau\rangle$ for HSPT or γ for KSPT is kept fixed as the continuum limit is approached. The case of LSPT is not considered explicitly as it is effectively covered by KSPT for $\gamma \to \infty$ or equivalently by a single-step HSPT algorithm.

Starting with HSPT, at the lowest perturbative order we expect autocorrelations to grow like L^2 when approaching the continuum limit if the average trajectory length $\langle\tau\rangle$ is kept fixed. On the other hand, the analysis of [41] shows how this scaling can be improved by choosing the average trajectory

length proportional to the correlation length of the system: $\langle\tau\rangle \propto 1/m \propto L$. Heuristically, the idea is that by adjusting the trajectory length with the correlation length one avoids the situation where configuration space is explored by a random walk, namely in random steps that are short compared with the natural scale of the system. What happens to the autocorrelations in HSPT beyond the tree-level dynamics, however, remains to be seen.

In Figure 6.5 we compare the results for the integrated autocorrelation $A_I(E_i)$ of the perturbative orders E_i as the continuum limit is approached. We compared the case where the average trajectory length was kept fixed at $\langle\tau\rangle = 1$ with the case where we set $\langle\tau\rangle = 1/m$ for the range of lattice sizes $4 \leq L \leq 32$. The step-size was adjusted so as to keep the errors in the equilibrium distribution roughly constant as L was increased, namely $\delta t = 2/L$ using the OMF4 integrator. We measured the observables after each trajectory, and chose $z = 4$ and $c = 0.2$. As can be seen from the figure the free field theory expectation also applies for the high-order fields: for the case where $\langle\tau\rangle = 1$ we observed the asymptotic random walk behaviour $A_I(E_i) \propto L^2$ whereas for $\langle\tau\rangle = 1/m$ the integrated autocorrelations were constant as the continuum limit was approached.

For KSPT the results from free field theory [41] indicate that at the lowest order in perturbation theory the autocorrelations are expected to increase as L^2 as the continuum limit is approached if the parameter γ is kept fixed. However, they increase only as L if $\gamma \propto m$ (see also [39]).[4] Hence γ effectively plays the role of an inverse trajectory length for the algorithm [36]. In Figure 6.6 we report the results for $A_I(E_i)$ for these two cases. In the first case we fixed $\gamma = 2$ as $L \to \infty$, while in the second case we set $\gamma = 2m$. Unlike the case of HSPT we chose a fixed step-size $\delta t = 0.25$, and we kept this constant as $L \to \infty$.[5] As we can see from the figure the two cases agree with the free field theory expectations for all the perturbative orders we investigated.

In conclusion, it seems that the free field theory expectations for autocorrela-

[4] We assume that the observables are measured at fixed stochastic time intervals as $L \to \infty$.

[5] We checked up to $L = 20$ that compatible results for the integrated autocorrelations were obtained if $\delta t \propto 1/L$ and the autocorrelations measured in units of this step-size were rescaled $\propto L$.

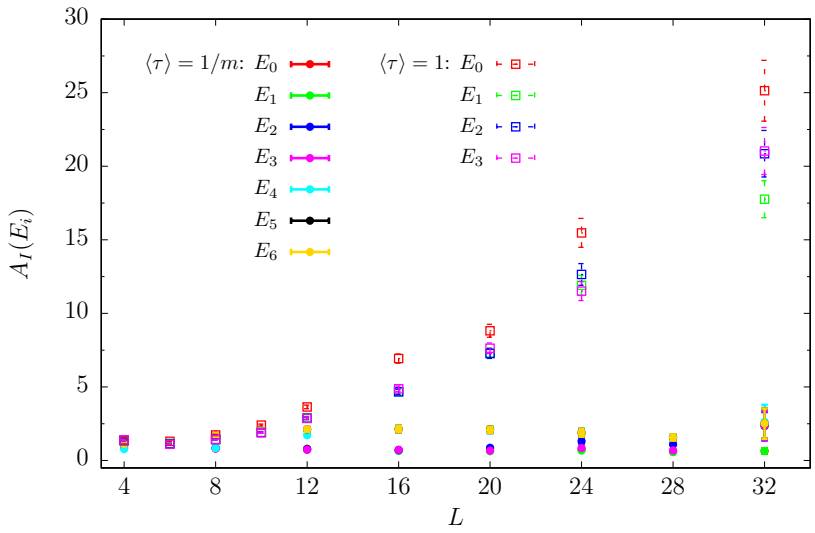


Figure 6.5 *Continuum scaling of the integrated autocorrelations $A_I(E_i)$ in HSPT for the cases $\langle\tau\rangle = 1$ and $\langle\tau\rangle = 1/m$. For $\langle\tau\rangle = 1$ we show results only up to $O(g_0^3)$, while for $\langle\tau\rangle = 1/m$ they go up to $O(g_0^6)$. The data is for $z = 4$ and $c = 0.2$. The errors on the integrated autocorrelations were estimated using the Γ-method [3].*

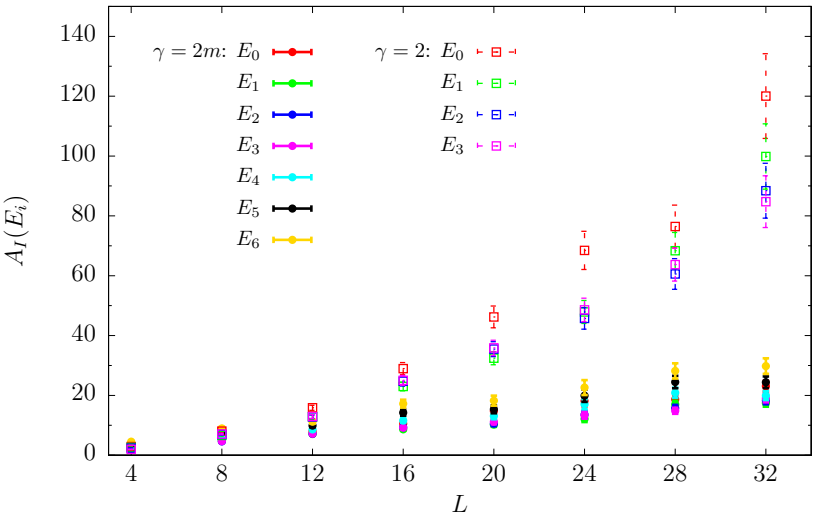


Figure 6.6 *Continuum scaling of the integrated autocorrelations $A_I(E_i)$ in KSPT for the cases $\gamma = 2$ and $\gamma = 2m$. For $\gamma = 2$ we show results only up to $O(g_0^3)$, while for $\gamma = 2m$ they go up to $O(g_0^6)$. The data is for $z = 4$, $c = 0.2$, and $\delta t = 0.25$. The errors on the integrated autocorrelations were estimated using the Γ-method [3].*

tions of the HMD and Kramers algorithms apply up to relatively high perturbative orders in the corresponding NSPT implementations.[6] Except for the case of KSPT at fixed γ this is a non-trivial result in view of the non-renormalizability of the HMD and SMD equations [26, 39].

6.0.4 Continuum variance scaling

Having investigated the dependence of the continuum scaling of the integrated autocorrelations for different algorithmic parameter scalings, we next studied the corresponding scaling of the variances $\mathrm{Var}(E_i)$. In Figure 6.7 we present results for the ratios $\mathrm{Var}(E_i)/\mathrm{Var}(E_0)$ with $i = 1, 2, 3$ for HSPT, comparing the cases $\langle\tau\rangle = 1$ and $\langle\tau\rangle = 1/m$ as $L \to \infty$. For convenience the results are normalized by their values at $L = 4$. As usual we chose $z = 4$, $c = 0.2$, and took $4 \leq L \leq 32$ and $\delta t = 2/L$. Recall that the lowest order variance $\mathrm{Var}(E_0)$ is independent on the algorithmic parameters, namely $\langle\tau\rangle$ (or γ below), and up to $O(1/L^2)$ corrections is constant with L. Observe that upon setting $\langle\tau\rangle = 1/m$ the variances $\mathrm{Var}(E_i)$ with $i > 1$ increase significantly as the continuum limit is approached. This effect is more pronounced as the perturbative order increases; on the other hand for $\langle\tau\rangle = 1$ the variances for all the perturbative orders considered grow very slowly with L and do not change significantly over the whole range of lattice sizes studied.

In Figure 6.8 we plot the results for the ratios of $\mathrm{Var}(E_i)/\mathrm{Var}(E_0)$ as obtained with KSPT. The two cases $\gamma = 2m$ and $\gamma = 2$ are shown. These results are very similar to those for HSPT: $\gamma = 2m$ leads to larger variances than keeping $\gamma = 2$ fixed, and these variances grow rapidly with perturbative order as the continuum limit is approached.

These results show that beyond the lowest perturbative order not only do the autocorrelations of observables computed using NSPT depend on the parameters of the algorithms but their variances do too. This is quite a different situation to the familiar case of non-perturbative computations.

[6] We also studied the dependence of the integrated autocorrelations $A_I(E_i)$ on the step-size δt and γ for KSPT, and on $\langle\tau\rangle$ for HSPT at fixed L and m. In this case the free field theory predictions of [41] also hold for all the perturbative orders we investigated.

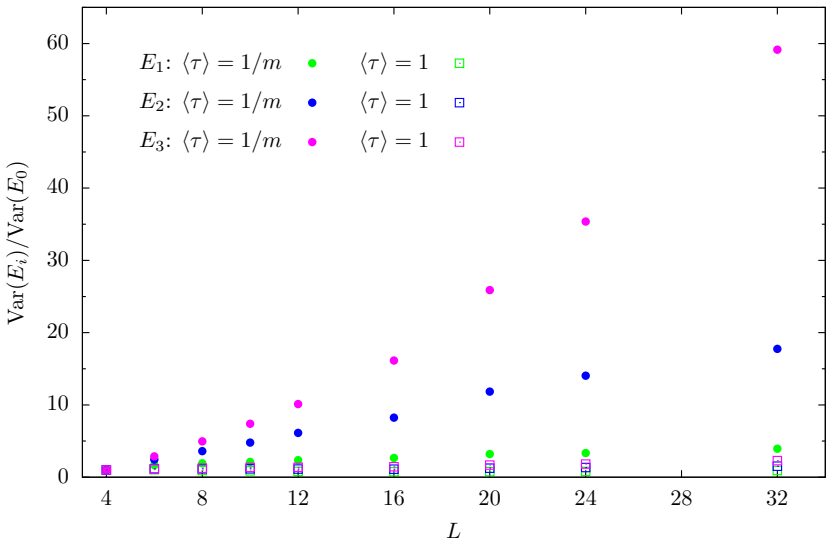


Figure 6.7 *Continuum scaling of the ratios* $\mathrm{Var}(E_i)/\mathrm{Var}(E_0)$ *with* $i = 1, 2, 3$ *for HSPT, for the cases* $\langle\tau\rangle = 1/m$ *and* $\langle\tau\rangle = 1$. *The case of* $z = 4$ *and* $c = 0.2$ *is shown, and the data are normalized at* $L = 4$.

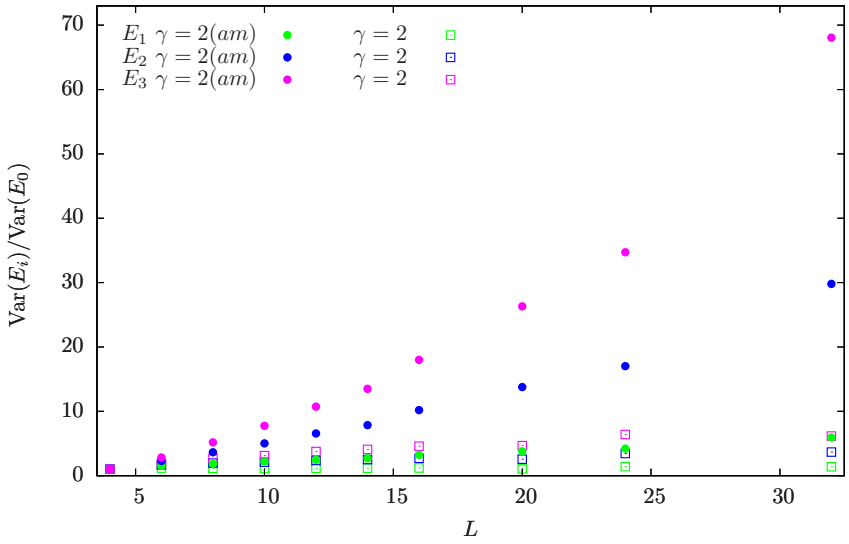


Figure 6.8 *Continuum scaling of the ratios* $\mathrm{Var}(E_i)/\mathrm{Var}(E_0)$ *for* $i = 1, 2, 3$ *for KSPT, for the cases* $\gamma = 2m$ *and* $\gamma = 2$.

6.0.5 Continuum cost scaling: the case of KSPT

From the results of the previous subsections it is clear that the most cost-effective tuning of parameters for an NSPT simulation is not trivial to determine. For all cases considered decreasing autocorrelations occurs concomitantly with increasing variances; the optimal compromise between the two effects must be found.

The situation is clear if we look directly at the total error (6.0.2) rather than at autocorrelations and variances separately, and compare the two parameter scalings investigated above. For illustration we consider the case of KSPT; HSPT gives very similar results. In Figure 6.9 we compare the relative error $\Delta\mathcal{E}_i/\mathcal{E}_i$ with $i = 0, 1, 2$ for the cases $\gamma = 2$ and $\gamma = 2m$. The number of configurations for the two parameter scalings is fixed to $N_{\text{config}} = 10^6$ for all the lattice sizes $4 \leq L \leq 20$. As usual the data is for $z = 4$, $c = 0.2$. We took $\delta t = 2/L$ and adjusted the measurement frequency $\propto L$. As expected, setting $\gamma = 2m$ is beneficial compared to having $\gamma = 2$ at the lowest perturbative order (top panel of Figure 6.9). On the other hand, when considering higher perturbative orders the case $\gamma = 2m$ seems to give comparable if not larger errors than fixing $\gamma = 2$ as $L \to \infty$. Hence, for the range of lattice sizes and perturbative orders we investigated, the effect of having smaller autocorrelations for $\gamma = 2m$ appears to be compensated if not overcome by the corresponding increase of the variances.

It appears clear that optimizing the performance of the algorithms requires finding the optimal value of $\langle\tau\rangle$ or γ for given lattice parameters, given observables, and the perturbative orders of interest. Focusing on the case of KSPT again, in Figure 6.10 we plot for example the relative errors $\Delta\mathcal{E}_i/\mathcal{E}_i$ for $i = 0, 1, 2$ as a function of γ for different values of L. For each L and perturbative order, the total number of configurations N_{config} was kept constant as γ was varied, and the results are normalized by their values at $\gamma = 2$. As usual $z = 4$, $c = 0.2$, and $\delta t = 2/L$. At tree-level (top panel) increasing γ leads to an increase of the relative error except at very small γ values and small lattice size. This is expected because in this case the variance is independent of γ, while the autocorrelations increase with γ until they saturate at some large enough value. In this regime the algorithm is effectively integrating the Langevin equation up to step-size errors.

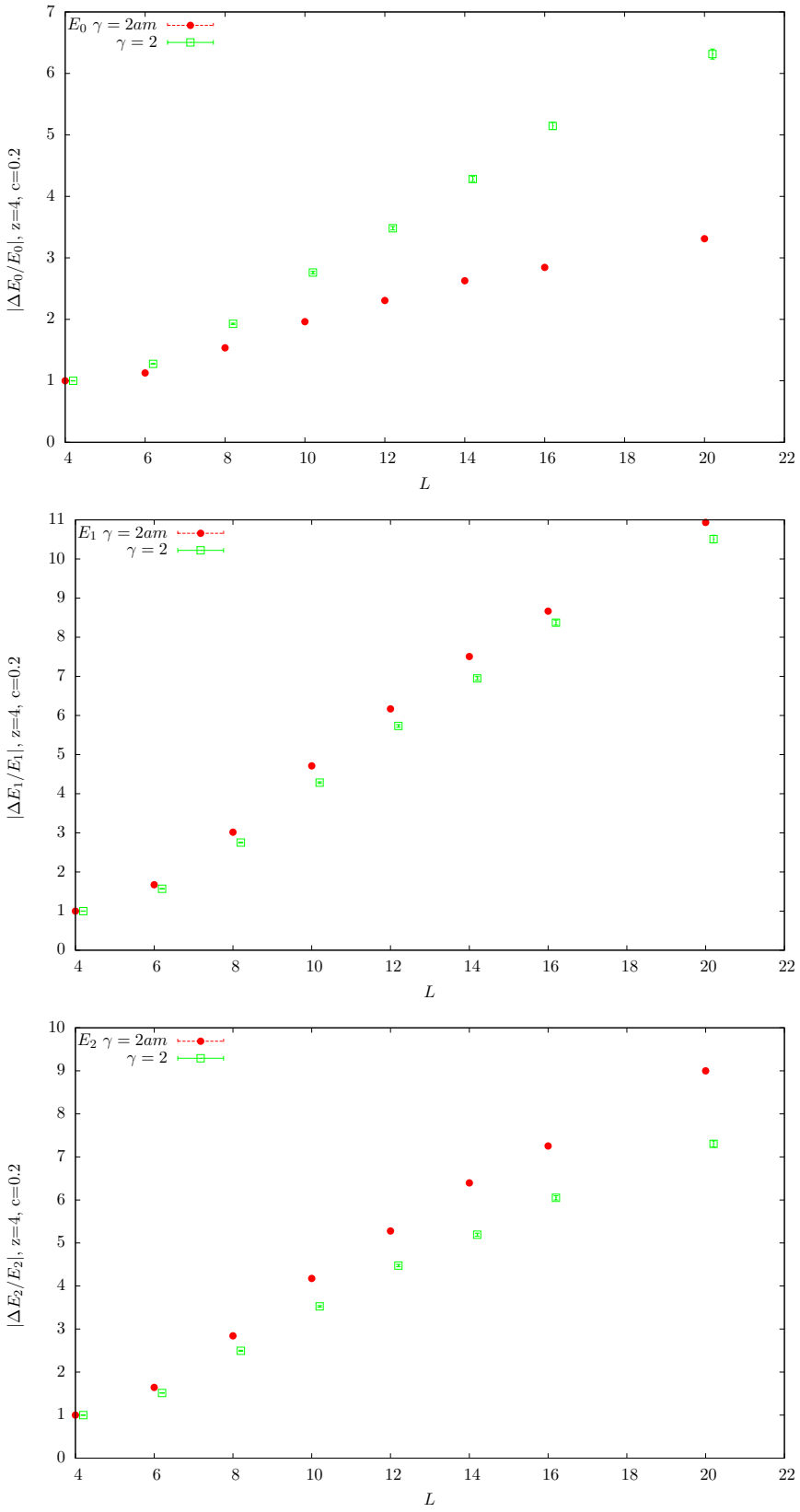


Figure 6.9 *Relative error $\Delta\mathcal{E}_i/\mathcal{E}_i$, $i = 0, 1, 2$, as a function of L for the two cases $\gamma = 2$ and $\gamma = 2m$. The errors in estimation of $\Delta\mathcal{E}_i/\mathcal{E}_i$ is smaller then the point size. $\mathcal{E}_i$ have been computed with LPT (Table 2.1) up machine precision, we didn't include these uncertainties in the error propagation. The data are normalized at $L = 4$.*

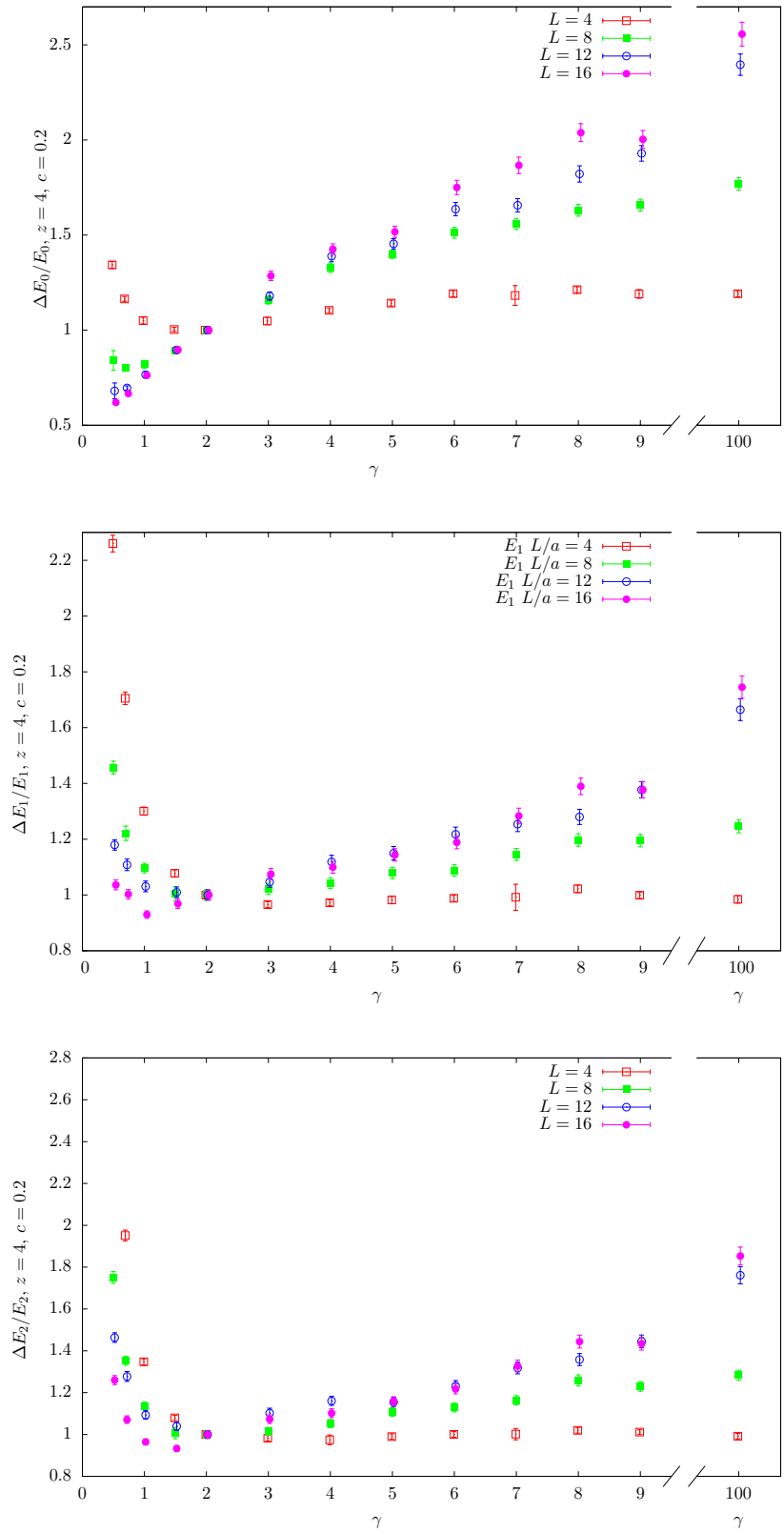


Figure 6.10 *Relative errors $\Delta\mathcal{E}_i/\mathcal{E}_i$ with $i = 0, 1, 2$ as a function of γ for $L = 4, 8, 12, 16$. $\mathcal{E}_i$ have been computed with LPT (Table 2.1) up machine precision, we didn't include these uncertainties in the error propagation. The data are normalized at $\gamma = 2$, and the results for $\gamma = 100$ are also shown. At this large value of γ the algorithm is effectively integrating the Langevin equation.*

The situation for the higher perturbative orders $\mathcal{E}_1$ and $\mathcal{E}_2$ is quite different. For small γ the errors fall rapidly as γ is increased; as we expect the autocorrelations to be small in this case we interpret this as a rapid fall of the variances. For larger γ values the errors increase only mildly compared to the situation at tree level. As at tree-level autocorrelations tend to grow with γ, but this effect is compensated by the variances decreasing as γ is increased. In particular, we note that the Langevin limit $\gamma \to \infty$ is characterized by having the largest autocorrelations but the smallest variances. There is a region of γ values for which the errors are minimized; in the example considered this does not appear to strongly depend on either the perturbative order or the lattice size. This is comforting as it allows us to tune γ easily and to improve the efficiency of the algorithm relative to Langevin.

Cost comparison with LSPT

The results of the previous subsection show that a proper tuning of the parameter γ increases the efficiency of KSPT over its Langevin limit, $\gamma \to \infty$. In the specific example considered, choosing a value of $\gamma \approx 2$ appears to be a good compromise for the different perturbative orders investigated, and it leads to a reduction of the statistical errors at fixed cost by a factor $\approx 1.5 - 2$ for $L = 16$ as compared to $\gamma \to \infty$; this corresponds to a factor $\approx 2 - 4$ in the cost at fixed statistical precision.

It is interesting to consider a direct comparison between KSPT and LSPT. We note that in practice LSPT differs from KSPT at $\gamma = \infty$ only by the different integration scheme used to integrate the Langevin equation. Hence, this comparison permits us to quantify the benefits of using efficient higher-order symplectic integrators in conjunction with a proper tuning of γ.

To this end, we compared the computational cost for computing the coefficients $\mathcal{E}_i$, $i = 0, \ldots 3$, to a specified statistical accuracy using KSPT with $\gamma = 2$ and LSPT. We chose to carry out this comparison with $L = 12$, $z = 4$ and $c = 0.2$. For $L = 12$ the reduction in the cost for a given statistical precision compared to $\gamma \to \infty$ is a factor $2 - 3$ for $\mathcal{E}_k$, $k = 1, 2, 3$, and a factor 6 for $\mathcal{E}_0$ (q.v., Figure 6.10).

In view of the results of §6.0.1, we chose $\delta t = 0.5$ for KSPT as we expect step-size errors to be very small compared to the precision of this test (see below). Similarly, for LSPT we took $\varepsilon = 0.01$, which corresponds to the smallest step-size considered in §6.0.1. At this value we also expect step-size errors to be small, and this is the most expensive of the simulations considered in the extrapolation $\varepsilon \to 0$. At each step we made measurements for both KSPT and LSPT, and considered a total of configurations $N = 4 \times 10^6$ and 4×10^7, respectively. The results are collected in Table 6.1.

	$\mathcal{E}_0 \times 10^5$	$\mathcal{E}_1 \times 10^8$	$\mathcal{E}_2 \times 10^9$	$\mathcal{E}_3 \times 10^{10}$
LSPT	2.2367(37)	−4.86(13)	2.352(54)	−1.599(40)
KSPT	2.2347(22)	−4.74(12)	2.223(49)	−1.517(35)
LPT	2.2347	−4.76	2.270	−

Table 6.1 *Results for $\mathcal{E}_i$, $i = 0, \ldots 3$ for $L = 12$, $z = 4$, $c = 0.2$ as obtained using KSPT with $\gamma = 2$ and LSPT. For KSPT we chose $\delta t = 0.5$ while for LSPT $\varepsilon = 0.01$, and measured at each step. The total number of configurations generated with the two algorithms is $N = 4 \times 10^6$ and 4×10^7 for KSPT and LSPT, respectively. The analytic perturbative results (LPT) for $\mathcal{E}_0$, $\mathcal{E}_1$, and $\mathcal{E}_2$ are also given for comparison.*

The KSPT and LSPT results are statistically consistent with each other and with the closed-form perturbative results (LPT) where these are available. KSPT and LSPT have approximately the same statistical errors with the numbers of configurations generated. The computer time used for updating a 12^4 lattice on a single core of an Intel Xeon E5-2630 Processor (2.4 GHz) is $0.21s$ for LSPT and $0.42s$ for KSPT: this is just the expected ratio of costs between the RK2 and OMF4 integrators, with the observation that this cost is dominated by the force computation.[7]

Thus, after rescaling the number of configurations N to have equal statistical errors, it becomes apparent that KSPT is $\approx 5 - 7$ times more cost effective than LSPT is in reaching a given statistical precision on the higher-order coefficients $\mathcal{E}_k$, $k = 1, 2, 3$, and roughly 14 times more cost effective for $\mathcal{E}_0$. As KSPT at fixed

[7] We recall that the RK2 integrator requires three force computations per step whereas the OMF4 integrator requires six (q.v. §3.4 and §4.2, respectively).

γ and LSPT have the same continuum scaling behavior in terms of variances and autocorrelations, one may expect a similar gain as $L \to \infty$. Furthermore, as the continuum limit is approached, it is advisable to reduce the step-size so as to keep systematic effects under control, and here again higher-order integrators are more cost effective.

Chapter 7

Conclusions

NSPT is a powerful technique that permits automation of high order perturbative computations on the lattice. As well as providing perturbative lattice estimates of quantities of interest these methods are interesting for extracting continuum perturbation theory results in cases where these are difficult or infeasible to obtain with continuum perturbative methods. However, to this end one needs efficient NSPT algorithms in order to be able to obtain precise results with both systematic and statistical errors under control. In particular, such results are desirable for a collection of lattice resolutions close to the continuum limit so that reliable continuum extrapolations may be performed.

In this work we investigated some new formulations of NSPT beyond LSPT, with the goal of finding more cost-effective algorithms. The new formulations of NSPT considered were ISPT (§5) and two particular cases of the GHMD (§4) algorithms; the HSPT and KSPT algorithms. All the algorithm have been discussed from a theoretical point of view in the chapter 3,4,5 and then studied numerically in 6.

Regarding ISPT [16] we can conclude that the first manifest advantage of this method over standard LSPT is that the results obtained are exact within statistical errors. Secondly, the stochastic field representing the theory to some

given order in the couplings is constructed directly from a set of Gaussian random fields, which are easy to generate. Despite these attractive features this algorithm has severe limitations beyond the lowest perturbative orders. First, similarly to conventional diagrammatic perturbation theory, the number of diagrams to be computed grows very rapidly with the perturbative order. While the cost of evaluating the diagrams is essentially proportional to the system size, their number increases exponentially as the perturbative order is increased. Most importantly, as shown by the present study , as the continuum limit is approached the statistical variance of perturbative coefficients computed using ISPT grows with increasing powers of L as the perturbative order is increased. Consequently it appears difficult to extract precise high-order results close to the continuum limit using this technique. While the exact details of our investigation certainly depend on the theory we considered, our conclusions are not specific to φ^4-theory. This has been confirmed by a recent study in the pure SU(3) Yang–Mills theory [50], where the nature of the divergences of the variances was also elucidated. In summary, the utility of this technique may be limited to a few low perturbative orders, which can nonetheless be of interest for some particularly difficult problems.

The other techniques investigated are two particular cases of the GHMD algorithms, HSPT and KSPT. The convergence of the GHMD algorithms have been established in §4.4. Respect to the Langevin implementation, HSPT and KSPT allow for a much more accurate discretization of the relevant equations. This is so because very efficient high-order symplectic integrators can be employed for the numerical integration of the MD equations. With such integrators the magnitude of the systematic errors is drastically reduced for a given number of force computations, and in practice one can run these algorithms with a small enough step-size that step-size extrapolations can be avoided.

As opposed to LSPT, HSPT and KSPT have tunable parameters, the average trajectory length $\langle\tau\rangle$ and the amount of partial momentum refreshment γ respectively, which may be adjusted so as to optimize their efficiency. However, beyond the lowest perturbative order finding the most cost-effective tuning of these parameters is not immediately obvious, in particular because their optimal continuum scaling is not trivial. The situation is complicated by the fact

that, unlike the more familiar non-perturbative simulations, not only do the autocorrelations of the perturbative coefficients computed in NSPT depend on the parameters of the chosen algorithm, but so do their variances. The general trend we observed is that when an algorithm is tuned to have small autocorrelations, the corresponding variances tend to increase, and therefore a trade-off between these two effects must be found. Moreover, only in the Langevin limit analytic understanding of the continuum scaling of both autocorrelations and variances is available, where the known result of [27] and [33] have been obtained from a different point of view in §3.3. The analysis indicates that the behavior of the autocorrelations of the high order fields with respect to the algorithmic parameters is the same as in the free field case. The behaviour of the variances is not easily predicted, and it seems to be different for different perturbative orders. A consequence of this is the fact that the optimal parameter scaling suggested by free field theory is not optimal when higher perturbative orders are considered. In our study we did not observe a significant difference in the cost with respect to the Langevin scaling of the algorithms (§6.0.5). Finding the optimal parameter scaling might thus be difficult, as it probably depends on the details of the calculation considered, i.e., the observables, the perturbative orders, and range of lattice sizes of interest. Nonetheless, when investigating the dependence of the errors in KSPT as a function of γ (§6.0.5) we found that for $\gamma \approx 2$ the algorithm is better than in its Langevin limit $\gamma \to \infty$, particularly so for large L. For example, at $L = 16$ an improvement by a factor of 1.5–2 was observed, depending on the order. This was possible since for the observables studied the optimal value of γ did not seem to depend strongly on either L or the perturbative order. Keeping this value fixed as $L \to \infty$ thus improves the efficiency of the algorithm over LSPT, although the scaling behaviour is the same. The question is whether this optimization is possible in general: this may need to be investigated on a case by case basis.

An additional outcome of this investigation is the fact that HSPT and KSPT appear to be in fact equivalent in terms of performances. For $\langle\tau\rangle \approx 1/\gamma$ the two algorithms have comparable autocorrelations and variances.

The novel NSPT methods presented here offer a simple and natural development from the standard Langevin-based algorithms. In particular, we heve

provided evidence that they can improve on previous methods hence allowing more precise results. These methods have been used and are under further development for the more interesting case of gauge theories [46, 51].

Appendix A

Renormalization procedure

A.1 Coupling renormalization

In regularized ϕ^4 theory we may compute an observable $\mathcal{O}$ as a perturbative expansion in the bare coupling g_0. However, in order to take the continuum limit of its expectation value, it is first of all necessary to express this perturbative series in terms of a renormalized coupling g. Of course, at finite lattice cutoff, the two are entirely equivalent as formal expansions and may readily be transformed into each other.

Suppose we have computed the perturbative expansion of the renormalized coupling g as a power series in the bare coupling g_0,

$$g = g_0 + \sum_{k \geq 2} c_k g_0^k. \tag{A.1.1}$$

We may then *revert* the expansion of g in terms of g_0 by writing (A.1.1) as

$$g_0 = g - \sum_{k \geq 2} c_k g_0^k, \tag{A.1.2}$$

and then recursively substituting (A.1.2) into itself to obtain

$$\begin{aligned} g_0 &= g - \sum_{k\geq 2} c_k \Big(g - \sum_{\ell \geq 2} c_\ell g_0^\ell\Big)^k \\ &= g - c_2 g^2 + (2c_2^2 - c_3) g^3 + (-5c_2^3 + 5c_2 c_3 - c_4) g^4 \\ &\quad + (14 c_2^4 - 21 c_2^2 c_3 + 6 c_2 c_4 + 3 c_3^2 - c_5) g^5 \\ &\quad + (-42 c_2^5 + 84 c_2^3 c_3 - 28 c_2^2 c_4 - 28 c_2 c_3^2 + 7 c_2 c_5 + 7 c_3 c_4 - c_6) g^6 + \cdots , \end{aligned} \tag{A.1.3}$$

noting that $O(g_0^N) = O(g^N)$.

Suppose that we have also computed the expansion of some operator of interest $\mathcal{O}$ in powers of g_0

$$\mathcal{O} = \sum_{k\geq 0} \mathcal{O}_k g_0^k, \tag{A.1.4}$$

Then by substituting (A.1.3) into (A.1.4) we obtain an expression for the expansion of $\mathcal{O}$ in powers of g:

$$\begin{aligned} \mathcal{O} = \mathcal{O}_0 &+ \mathcal{O}_1 g + (-c_2 \mathcal{O}_1 + \mathcal{O}_2) g^2 + \big((2c_2^2 - c_3)\mathcal{O}_1 - 2c_2 \mathcal{O}_2 + \mathcal{O}_3\big) g^3 \\ &+ \big((-5c_2^3 + 5c_2 c_3 - c_4)\mathcal{O}_1 + (5c_2^2 - 2c_3)\mathcal{O}_2 - 3c_2 \mathcal{O}_3 + \mathcal{O}_4\big) g^4 \\ &+ \big((14 c_2^4 - 21 c_2^2 c_3 + 6 c_2 c_4 + 3 c_3^2 - c_5)\mathcal{O}_1 + (-14 c_2^3 + 12 c_2 c_3 - 2 c_4)\mathcal{O}_2 \\ &\qquad + (9c_2^2 - 3c_3)\mathcal{O}_3 - 4c_2 \mathcal{O}_4 + \mathcal{O}_5\big) g^5 \\ &+ \big((-42 c_2^5 + 84 c_2^3 c_3 - 28 c_2^2 c_4 - 28 c_2 c_3^2 + 7 c_2 c_5 + 7 c_3 c_4 - c_6)\mathcal{O}_1 \\ &\qquad + (42 c_2^4 - 56 c_2^2 c_3 + 14 c_2 c_4 + 7 c_3^2 - 2 c_5)\mathcal{O}_2 \\ &\qquad + (-28 c_2^3 + 21 c_2 c_3 - 3 c_4)\mathcal{O}_3 + (14 c_2^2 - 4 c_3)\mathcal{O}_4 - 5 c_2 \mathcal{O}_5 + \mathcal{O}_6\big) g^6 + \cdots \end{aligned}$$

For the numerical computation of the perturbative expansion of $\mathcal{O}$ we are therefore free to consider an expansion in powers of g_0 as this is entirely equivalent — as formal power series — to expansion in powers of g.

A.2 Mass renormalization

The stochastic field ϕ is considered to be of the form

$$\phi(x) = \sum_{k,l\geq 0} \phi_{k,\ell}(x) g_0^k (\delta m^2)^\ell \tag{A.2.1}$$

where g_0 is the bare coupling and δm^2 is the mass counterterm.[1] Once the table of numbers $\phi_{k,\ell}$ has been computed, the expectation value $\langle \cdots \rangle_\eta$ of functions of these quantities may be estimated, but they must be fitted to the renormalization conditions in order to compute physical quantities. Here we shall present algebraic expressions for the formal power series manipulation in order to explain the renormalization procedure; in actual computations we automated these formal manipulations using the numerical values of the coefficients.

The renormalization condition (2.1.1) that defines m^2 can be rewritten as

$$m^2 = \hat{p}_*^2 \frac{\chi_2^*}{\chi_2 - \chi_2^*}.$$

Therefore, since we can calculate χ_2 and χ_2^* as power series in both g_0 and δm^2

$$\begin{aligned}
\chi_2 = {} & \Big\langle \tilde{\phi}_{0,0}(0)^2 \Big\rangle + 2\Big\langle \tilde{\phi}_{0,0}(0)\tilde{\phi}_{0,1}(0) \Big\rangle \delta m^2 + \Big\langle 2\tilde{\phi}_{0,0}(0)\tilde{\phi}_{0,2}(0) + \tilde{\phi}_{0,1}(0)^2 \Big\rangle \delta m^4 \\
& + 2\Big\langle \tilde{\phi}_{0,0}(0)\tilde{\phi}_{0,3}(0) + \tilde{\phi}_{0,1}(0)\tilde{\phi}_{0,2}(0) \Big\rangle \delta m^6 \\
& + 2\Big(\Big\langle \tilde{\phi}_{0,0}(0)\tilde{\phi}_{1,0}(0) \Big\rangle + \Big\langle \tilde{\phi}_{0,0}(0)\tilde{\phi}_{1,1}(0) + \tilde{\phi}_{1,0}(0)\tilde{\phi}_{0,1}(0) \Big\rangle \delta m^2 \\
& \quad + \Big\langle \tilde{\phi}_{0,0}(0)\tilde{\phi}_{1,2}(0) + \tilde{\phi}_{1,0}(0)\tilde{\phi}_{0,2}(0) + \tilde{\phi}_{0,1}(0)\tilde{\phi}_{1,1}(0) \Big\rangle \delta m^4 \Big) g_0 \\
& + \Big(\Big\langle 2\tilde{\phi}_{0,0}(0)\tilde{\phi}_{2,0}(0) + \tilde{\phi}_{1,0}(0)^2 \Big\rangle \\
& \quad + 2\Big\langle \tilde{\phi}_{0,0}(0)\tilde{\phi}_{2,1}(0) + \tilde{\phi}_{1,0}(0)\tilde{\phi}_{1,1}(0) + \tilde{\phi}_{0,1}(0)\tilde{\phi}_{2,0}(0) \Big\rangle \delta m^2 \Big) g_0^2 \\
& + 2\Big\langle \tilde{\phi}_{0,0}(0)\tilde{\phi}_{3,0}(0) + \tilde{\phi}_{1,0}(0)\tilde{\phi}_{2,0}(0) \Big\rangle g_0^3 + O(g_0^4)
\end{aligned}$$

[1] Remember that δm^2 has contributions of order g_0^n for $n \geq 1$ when it has been determined from the renormalization conditions (see the following discussion).

$$
\begin{aligned}
\chi_2^* = &\left\langle \tilde\phi_{0,0}(p_*)\tilde\phi_{0,0}(-p_*)\right\rangle + \left\langle \tilde\phi_{0,0}(p_*)\tilde\phi_{0,1}(-p_*) + \tilde\phi_{0,0}(-p_*)\tilde\phi_{0,1}(p_*)\right\rangle \delta m^2 \\
&+ \left\langle \tilde\phi_{0,0}(p_*)\tilde\phi_{0,2}(-p_*) + \tilde\phi_{0,1}(p_*)\tilde\phi_{0,1}(-p_*) + \tilde\phi_{0,0}(-p_*)\tilde\phi_{0,2}(p_*)\right\rangle \delta m^4 \\
&+ \Big\langle \tilde\phi_{0,0}(p_*)\tilde\phi_{0,3}(-p_*) + \tilde\phi_{0,1}(p_*)\tilde\phi_{0,2}(-p_*) \\
&\quad + \tilde\phi_{0,2}(p_*)\tilde\phi_{0,1}(-p_*) + \tilde\phi_{0,0}(-p_*)\tilde\phi_{0,3}(p_*)\Big\rangle \delta m^6 \\
&+ \Big(\left\langle \tilde\phi_{0,0}(p_*)\tilde\phi_{1,0}(-p_*) + \tilde\phi_{0,0}(-p_*)\tilde\phi_{1,0}(p_*)\right\rangle \\
&\quad + \left\langle \tilde\phi_{0,0}(p_*)\tilde\phi_{1,1}(-p_*) + \tilde\phi_{1,0}(p_*)\tilde\phi_{0,1}(-p_*) \right. \\
&\qquad \left. + \tilde\phi_{0,1}(p_*)\tilde\phi_{1,0}(-p_*) + \tilde\phi_{0,0}(-p_*)\tilde\phi_{1,1}(p_*)\right\rangle \delta m^2 \\
&\quad + \Big\langle \tilde\phi_{0,0}(p_*)\tilde\phi_{1,2}(-p_*) + \tilde\phi_{1,0}(p_*)\tilde\phi_{0,2}(-p_*) \\
&\qquad + \tilde\phi_{0,1}(p_*)\tilde\phi_{1,1}(-p_*) + \tilde\phi_{1,1}(p_*)\tilde\phi_{0,1}(-p_*) \\
&\qquad + \tilde\phi_{0,2}(p_*)\tilde\phi_{1,0}(-p_*) + \tilde\phi_{0,0}(-p_*)\tilde\phi_{1,2}(p_*)\Big\rangle \delta m^4\Big) g_0 \\
&+ \Big(\left\langle \tilde\phi_{0,0}(p_*)\tilde\phi_{2,0}(-p_*) + \tilde\phi_{1,0}(p_*)\tilde\phi_{1,0}(-p_*) + \tilde\phi_{0,0}(-p_*)\tilde\phi_{2,0}(p_*)\right\rangle \\
&\quad + \left\langle \tilde\phi_{0,0}(p_*)\tilde\phi_{2,1}(-p_*) + \tilde\phi_{1,0}(p_*)\tilde\phi_{1,1}(-p_*) \right. \\
&\qquad \left. + \tilde\phi_{0,1}(p_*)\tilde\phi_{2,0}(-p_*) + \tilde\phi_{2,0}(p_*)\tilde\phi_{0,1}(-p_*) \right. \\
&\qquad \left. + \tilde\phi_{1,1}(p_*)\tilde\phi_{1,0}(-p_*) + \tilde\phi_{0,0}(-p_*)\tilde\phi_{2,1}(p_*)\right\rangle \delta m^2\Big) g_0^2 \\
&+ \Big(\Big\langle \tilde\phi_{0,0}(p_*)\tilde\phi_{3,0}(-p_*) + \tilde\phi_{1,0}(p_*)\tilde\phi_{2,0}(-p_*) \\
&\quad + \tilde\phi_{2,0}(p_*)\tilde\phi_{1,0}(-p_*) + \tilde\phi_{0,0}(-p_*)\tilde\phi_{3,0}(p_*)\Big\rangle\Big) g_0^3 + O(g_0^4)
\end{aligned}
$$

where we defined the Fourier transform of the coefficient fields as,

$$
\tilde\phi_{k,l}(p) = \frac{1}{L^2}\sum_{x\in\Omega} e^{-ipx}\phi_{k,l}(x), \quad p \in \tilde\Omega,
$$

we can multiply and invert[2] χ and χ^* to compute m^2 as power series in g_0 and δm^2

[2]The *inverse* of a power series $S(g_0, \delta m)$ the power series for $1/S(g_0, \delta m)$.

$$m^2 = \sum_{k,\ell \geq 0} a_{k,\ell}\, g_0^k (\delta m^2)^\ell \tag{A.2.2}$$

where the coefficients $a_{k,\ell}$ are:

$$a_{0,0} = \frac{\Big\langle \tilde\phi_{0,0}(p_*)\tilde\phi_{0,0}(-p_*)\Big\rangle}{\Big\langle \tilde\phi_{0,0}(0)^2 - \tilde\phi_{0,0}(p_*)\tilde\phi_{0,0}(-p_*)\Big\rangle}\,\hat p_*^2,$$

$$a_{1,0} = -\frac{\begin{array}{r} 2\Big\langle \tilde\phi_{0,0}(0)\tilde\phi_{1,0}(0)\Big\rangle\Big\langle \tilde\phi_{0,0}(p_*)\tilde\phi_{0,0}(-p_*)\Big\rangle \\ -\Big\langle \tilde\phi_{0,0}(0)^2\Big\rangle\Big\langle \tilde\phi_{0,0}(p_*)\tilde\phi_{1,0}(-p_*) + \tilde\phi_{0,0}(-p_*)\tilde\phi_{1,0}(p_*)\Big\rangle \end{array}}{\Big\langle \tilde\phi_{0,0}(0)^2 - \tilde\phi_{0,0}(p_*)\tilde\phi_{0,0}(-p_*)\Big\rangle^2}\,\hat p_*^2,$$

$$a_{0,1} = -\frac{\begin{array}{r} 2\Big\langle \tilde\phi_{0,0}(0)\tilde\phi_{0,1}(0)\Big\rangle\Big\langle \tilde\phi_{0,0}(p_*)\tilde\phi_{0,0}(-p_*)\Big\rangle \\ -\Big\langle \tilde\phi_{0,0}(0)^2\Big\rangle\Big\langle \tilde\phi_{0,0}(p_*)\tilde\phi_{0,1}(-p_*) + \tilde\phi_{0,0}(-p_*)\tilde\phi_{0,1}(p_*)\Big\rangle \end{array}}{\Big\langle \tilde\phi_{0,0}(0)^2 - \tilde\phi_{0,0}(p_*)\tilde\phi_{0,0}(-p_*)\Big\rangle^2}\,\hat p_*^2,$$

and so forth. By construction $a_{0,0} = m^2$, so at lowest order in g_0 the mass m is the mass that enters the scalar propagator (5.0.10). Having determined the coefficient $a_{k,\ell}$ in (A.2.2) we can now determine the coefficients m_k^2 of the expansion

$$\delta m^2 = \sum_{k\geq 1} m_k^2\, g_0^k$$

by imposing the relation (A.2.2) order by order in g_0, thus obtaining

$$
\begin{aligned}
m_1^2 &= -\frac{a_{1,0}}{a_{0,1}}, \\
m_2^2 &= -\frac{a_{2,0} + a_{1,1}m_1^2 + a_{0,2}m_1^4}{a_{0,1}}, \\
m_3^2 &= -\frac{a_{3,0} + a_{2,1}m_1^2 + 2a_{0,2}m_1^2m_2^2 + a_{1,2}m_1^4 + a_{1,1}m_2^2 + a_{0,3}m_1^6}{a_{0,1}}, \\
m_4^2 &= -\frac{\begin{array}{r} a_{4,0} + a_{3,1}m_1^2 + a_{2,1}m_2^2 + a_{1,1}m_3^2 + a_{2,2}m_1^4 + 2a_{1,2}m_1^2m_2^2 \\ +a_{0,2}m_2^4 + 2a_{0,2}m_1^2m_3^2 + a_{1,3}m_1^6 + 3a_{0,3}m_1^4m_2^2 + a_{0,4}m_1^8 \end{array}}{a_{0,1}}, \\
&\vdots
\end{aligned}
$$

Once δm^2 is determined, the field ϕ and any other observable previously computed as a series in δm^2 and g_0 can be reduced to a series in g_0 alone.

A.3 Wavefunction renormalization

The renormalization of a generic correlation function by the wavefunction renormalization, or any multiplicative renormalization factor, does not present any additional difficulty. We may compute Z as a power series in g_0 and δm^2 from the renormalization condition (2.1.2),

$$Z = m^2\chi = \sum_{k,\ell\geq 0} Z_{k,\ell}\, g_0^k (\delta m^2)^\ell,$$

where

$$Z_{0,0} = a_{0,0}\Big\langle \tilde\phi_{0,0}(0)^2 \Big\rangle$$

$$Z_{1,0} = a_{1,0}\Big\langle \tilde\phi_{0,0}(0)^2 \Big\rangle + 2a_{0,0}\Big\langle \tilde\phi_{1,0}(0)\tilde\phi_{0,0}(0) \Big\rangle$$

$$Z_{0,1} = a_{0,1}\Big\langle \tilde\phi_{0,0}(0)^2 \Big\rangle + 2a_{0,0}\Big\langle \tilde\phi_{0,0}(0)\tilde\phi_{0,1}(0) \Big\rangle$$

$$\begin{aligned} Z_{2,0} = {} & 2a_{0,0}\Big\langle \tilde\phi_{0,0}(0)\tilde\phi_{2,0}(0) \Big\rangle + a_{0,0}\Big\langle \tilde\phi_{1,0}(0)^2) \Big\rangle \\ & + 2a_{1,0}\Big\langle \tilde\phi_{0,0}(0)\tilde\phi_{1,0}(0) \Big\rangle + a_{2,0}\Big\langle \tilde\phi_{0,0}(0)^2) \Big\rangle \end{aligned}$$

$$\begin{aligned} Z_{1,1} = {} & 2a_{0,0}\Big\langle \tilde\phi_{0,0}(0)\tilde\phi_{1,1}(0) \Big\rangle + 2a_{0,0}\Big\langle \tilde\phi_{1,0}(0)\tilde\phi_{0,1}(0) \Big\rangle \\ & + 2a_{0,1}\Big\langle \tilde\phi_{0,0}(0)\tilde\phi_{1,0}(0) \Big\rangle + 2a_{1,0}\Big\langle \tilde\phi_{0,0}(0)\tilde\phi_{0,1}(0) \Big\rangle + a_{1,1}\Big\langle \tilde\phi_{0,0}(0)^2 \Big\rangle \end{aligned}$$

$$\begin{aligned} Z_{0,2} = {} & 2a_{0,0}\Big\langle \tilde\phi_{0,0}(0)\tilde\phi_{0,2}(0) \Big\rangle + a_{0,0}\Big\langle \tilde\phi_{0,1}(0)^2 \Big\rangle \\ & + 2a_{0,1}\Big\langle \tilde\phi_{0,0}(0)\tilde\phi_{0,1}(0) \Big\rangle + a_{0,2}\Big\langle \tilde\phi_{0,0}(0)^2 \Big\rangle, \end{aligned}$$

and so forth. We can now compute a renormalized correlation function as a power series in g_0 and the renormalized mass m as,

$$\begin{aligned} Z^{n/2}\big\langle \phi(x_1)\cdots\phi(x_n)\big\rangle = {} & \Big(\sum_{k,\ell\ge 0} Z_{k,\ell}\, g_0^k (\delta m^2)^\ell \Big)^{n/2} \\ & \times \Big\langle \sum_{k_1,\ell_1\ge 0} \phi_{k_1,\ell_1}(x_1) g_0^{k_1} (\delta m^2)^{\ell_1} \;\cdots \sum_{k_n,\ell_n\ge 0} \phi_{k_n,\ell_n}(x_n) g_0^{k_n} (\delta m^2)^{\ell_n} \Big\rangle \\ = {} & \Big(\sum_{k,l\ge 0} Z_{k,\ell}\, g_0^k \Big(\sum_{j\ge 1} m_j^2\, g_0^j \Big)^\ell \Big)^{n/2} \\ & \times \Big\langle \sum_{k_1,\ell_1\ge 0} \phi_{k_1,\ell_1}(x_1) g_0^{k_1} \Big(\sum_{j_1\ge 1} m_{j_1}^2\, g_0^{j_1} \Big)^{\ell_1} \cdots \sum_{k_n,\ell_n\ge 0} \phi_{k_n,\ell_n}(x_n) g_0^{k_n} \Big(\sum_{j_n\ge 1} m_{j_n}^2\, g_0^{j_n} \Big)^{\ell_n} \Big\rangle . \end{aligned}$$

This expansion in the bare coupling g_0 can be replaced by one in the renormalized coupling g as explained in §A.1, and the correlation function is then properly renormalized.

Appendix B

Errors and Autocorrelations

Here we will describe the technique we used to estimate the error in the Monte Carlo simulations. We implement the so called Γ method discussed in [3] and partially in the appendix of [52]. Let us assume that we have a set of primary observable labeled x_α with $\alpha = 1, 2, ...$ distributed with probability distribution $P[x_1, x_2, ...]$. The expectation value of x_α is defined as

$$\langle x_\alpha \rangle \equiv X_\alpha \equiv \int \prod_\beta dx_b P[x_1, x_2, ...] x_\alpha. \tag{B.0.1}$$

Our goal is to estimate a function of the expectation values

$$F \equiv f(X_1, X_2, ...). \tag{B.0.2}$$

B.1 Primary observable

To estimate it we use Monte Carlo estimate of the primary observable x_α^{ir} where $i = 1, ..., N_r$ count the successive estimate from the Markov update procedure and $r = 1, ..., R$ count the number of statistically independent replicas, the total

number of configuration is $N = RN_r$. We define mean within each replica

$$\bar{x}^r_\alpha = \frac{1}{N_r}\sum_{i=1}^{N_r} x^{ir}_\alpha \tag{B.1.1}$$

and the mean over all the replicas as

$$\bar{\bar{x}}_\alpha = \frac{1}{N}\sum_{r=1}^{R} N_r \bar{x}^i_\alpha, \tag{B.1.2}$$

The law of large number states that for large N_r $\bar{x}^r_\alpha$ and $\bar{\bar{x}}_\alpha$ goes to X_α and the central limit theorem say that their distribution is Gaussian. The two estimator $\bar{x}^r_\alpha$ and $\bar{\bar{x}}_\alpha$ are unbiased, i.e.

$$\langle \bar{\bar{x}}_\alpha - X_\alpha \rangle = \langle \bar{x}^r_\alpha - X_\alpha \rangle = 0. \tag{B.1.3}$$

The covariance Matrix of $\bar{x}_\alpha$ is

$$\left\langle (\bar{x}^r_\alpha - X_\alpha)(\bar{x}^s_\beta - X_\beta) \right\rangle = \frac{1}{N_r N_s}\sum_{i=1}^{N_r}\sum_{j=1}^{N_s} \left\langle (x^{i,r}_\alpha - X_\alpha)(x^{j,s}_\beta - X_\beta) \right\rangle . \tag{B.1.4}$$

If we assume that the replicas r and s are independent we have

$$\left\langle (x^{i,r}_\alpha - X_\alpha)(x^{j,s}_\beta - X_\beta) \right\rangle = \delta_{r,s}\Gamma_{\alpha\beta}(j-i) \tag{B.1.5}$$

where $\Gamma_{\alpha\beta}(j-i)$ is the autocorrelation function that depends only on the distance $j - i$. Substituting the (B.1.5) inside the (B.1.4) we get

$$\left\langle (\bar{x}^r_\alpha - X_\alpha)(\bar{x}^s_\beta - X_\beta) \right\rangle = \frac{1}{N_r^2}\sum_{i=1}^{N_r}\sum_{j=1}^{N_r} \delta_{r,s}\Gamma_{\alpha\beta}(j-i), \tag{B.1.6}$$

that setting $t = j - i$ we have

$$\left\langle (\bar{x}^r_\alpha - X_\alpha)(\bar{x}^s_\beta - X_\beta) \right\rangle = \frac{1}{N_r^2}\sum_{t=1-N_r}^{N_r-1}\sum_{j=1}^{N_r-t} \Gamma_{\alpha\beta}(t) = \frac{1}{N_r^2}\sum_{t=1-N_r}^{1+N_r} (N_r - t)\Gamma_{\alpha\beta}(t). \tag{B.1.7}$$

Assuming that $\Gamma_{\alpha\beta}(t) \sim e^{-t/\tau}$ for large t and $N_r >> \tau$ we can approximate

$$\sum_{t=1-N_r}^{1+N_r}(1-\frac{t}{N_r})\Gamma_{\alpha\beta}(t) = \sum_{t=-\infty}^{\infty}\Gamma_{\alpha\beta}(t) + O\left(\frac{\tau}{N_r}\right) = C_{\alpha\beta} + O\left(\frac{\tau}{N_r}\right) \tag{B.1.8}$$

where we define

$$C_{\alpha\beta} = \sum_{t=-\infty}^{\infty}\Gamma_{\alpha\beta}(t). \tag{B.1.9}$$

Thus for the covariance matrix of $\bar{x}^r_\alpha$ we have

$$\left\langle(\bar{x}^r_\alpha - X_\alpha)(\bar{x}^s_\beta - X_\beta)\right\rangle = \frac{1}{N_r}\delta_{r,s}C_{\alpha\beta} + O\left(\frac{\tau}{N_r}\right) \tag{B.1.10}$$

Similar for the covariance of matrix of $\bar{\bar{x}}_\alpha$

$$\begin{aligned}\langle(\bar{\bar{x}}_\alpha - X_\alpha)(\bar{\bar{x}}_\beta - X_\beta)\rangle &= \frac{1}{N^2}\sum_{r=1}^{R}\sum_{s=1}^{R}N_rN_s\left\langle(\bar{x}^r_\alpha - X_\alpha)(\bar{x}^s_\beta - X_\beta)\right\rangle \\ &= \frac{1}{N^2}\sum_{r=1}^{R}\left(N_rC_{\alpha\beta} + O\left(\frac{\tau}{N_r}\right)\right) = \frac{1}{N}C_{\alpha\beta} + O\left(\frac{R\tau}{N_r}\right),\end{aligned} \tag{B.1.11}$$

which tell us that $\bar{\bar{x}}_\alpha$ differs from X_α by a factor $\sqrt{C_{\alpha\beta}/N}$.

B.2 Derived quantities

For a function F of the primary observable x_α we considered two estimator

$$\bar{\bar{F}} = f(\bar{\bar{x}}_\alpha) \tag{B.2.1}$$

and

$$\bar{F} = \frac{1}{N}\sum_{r=1}^{R}N_rf(\bar{x}^r_\alpha). \tag{B.2.2}$$

Here we will assume that we can truncate the Taylor series around the exact value X_α of $\bar{\bar{F}}$ up to the second order

$$\bar{\bar{F}} = F + \sum_\alpha f_\alpha(\bar{\bar{x}}_\alpha - X_\alpha) + \frac{1}{2}\sum_{\alpha\beta} f_{\alpha\beta}(\bar{\bar{x}}_\alpha - X_\alpha)(\bar{\bar{x}}_\beta - X_\beta) + ... \tag{B.2.3}$$

where F_α and $F_{\alpha\beta}$ are the derivative computed in the exact value X_α

$$F_\alpha = \left.\frac{\partial f}{\partial x_\alpha}\right|_{x_\alpha = X_\alpha}, \qquad F_{\alpha\beta} = \left.\frac{\partial^2 f}{\partial x_\alpha \partial x_\beta}\right|_{x_{\alpha,\beta} = X_{\alpha,\beta}}. \tag{B.2.4}$$

We can compute the leading order bias of $\bar{\bar{F}}$ computing

$$\begin{aligned}\left\langle \bar{\bar{F}} \right\rangle &= F + \frac{1}{2}\sum_{\alpha\beta} f_{\alpha\beta} \left\langle (\bar{\bar{x}}_\alpha - X_\alpha)(\bar{\bar{x}}_\beta - X_\beta) \right\rangle + ... \\ &\approx F + \frac{1}{2N}\sum_{\alpha\beta} f_{\alpha\beta} C_{\alpha\beta}.\end{aligned} \tag{B.2.5}$$

If N is big enough this bias is negligible respect to the statistical errors.
The Taylor series of the estimator $\bar{F}$ is

$$\bar{F} = F + \frac{1}{N}\sum_{r=1}^{R} N_r \sum_\alpha f_\alpha(\bar{x}^r_\alpha - X_\alpha) + \frac{1}{2N}\sum_{r=1}^{R} N_r \sum_{\alpha\beta} f_{\alpha\beta}(\bar{x}^r_\alpha - X_\alpha)(\bar{x}^r_\beta - X_\beta) + ... \tag{B.2.6}$$

So for its expectation value we have

$$\begin{aligned}\left\langle \bar{F} \right\rangle &= F + \frac{1}{2N}\sum_{r}^{R} N_r \sum_{\alpha\beta} f_{\alpha\beta} \left\langle (\bar{x}^r_\alpha - X_\alpha)(\bar{x}^r_\beta - X_\beta) \right\rangle + ... \\ &\approx F + \frac{1}{2N}\sum_{\alpha\beta}\sum_{r=1}^{R} f_{\alpha\beta} C_{\alpha\beta} = F + \frac{R}{2N}\sum_{\alpha\beta} f_{\alpha\beta} C_{\alpha\beta}.\end{aligned} \tag{B.2.7}$$

If $R > 1$ it is possible combine the (B.2.5) and the (B.2.7) to eliminate the leading bias

$$\langle \mathcal{F} \rangle = \left\langle \frac{R\bar{\bar{F}} - \bar{F}}{R-1} \right\rangle \approx F. \tag{B.2.8}$$

Note that the Taylor series of $\bar{F}$ and $\bar{\bar{F}}$ are equal up to the first order (it can be verified plugging the (B.2.5) inside the (B.2.3)). This implies that the two estimator $\bar{F}$ and $\bar{\bar{F}}$ have the same leading error

$$\Delta^2_{\bar{F}} = \left\langle (\bar{F} - F)^2 \right\rangle \approx \left\langle (\bar{\bar{F}} - F)^2 \right\rangle = \Delta^2_{\bar{\bar{F}}} \tag{B.2.9}$$

The reason become clear computing explicitly one of the two

$$\begin{aligned}
\Delta^2_{\bar{F}} \approx \left\langle (\bar{F} - F)^2 \right\rangle &= \left\langle \left(F + \frac{1}{N} \sum_{r=1}^{R} N_r \sum_{\alpha} f_\alpha (\bar{x}^r_\alpha - X_\alpha) - F \right)^2 \right\rangle \\
&= \frac{1}{N^2} \sum_{r,s=1}^{R} N_r N_s \sum_{\alpha\beta} f_\alpha f_\beta \left\langle (\bar{x}^r_\alpha - X_\alpha)(\bar{x}^s_\beta - X_\beta) \right\rangle \\
&= \frac{1}{N} \sum_{\alpha\beta} f_\alpha f_\beta C_{\alpha\beta}.
\end{aligned} \tag{B.2.10}$$

Where in the first line we expand in Taylor (B.2.6) series up to the first order , in the second line we used the (B.1.11). I can be verified that also the Taylor expansion of $\mathcal{F}$ is equal to the Taylor series of $\bar{F}$ up to the first order so the leading error of the estimator $\mathcal{F}$ will be equal to $\Delta_{\bar{F}}$

$$\Delta^2_{\mathcal{F}} = \left\langle (\mathcal{F} - F)^2 \right\rangle = \left\langle \left(\frac{R\bar{\bar{F}} - \bar{F}}{R-1} - F \right)^2 \right\rangle \tag{B.2.11}$$

$$\approx \left\langle \left(F + \frac{1}{N} \sum_{r=1}^{R} N_r \sum_{\alpha} f_\alpha (\bar{x}^r_\alpha - X_\alpha) - F \right)^2 \right\rangle = \Delta^2_{\bar{F}}. \tag{B.2.12}$$

We may rewrite the expression for the error of $\mathcal{F}$ as

$$\Delta^2_{\mathcal{F}} = \frac{2A_{I,F}\mathrm{Var}_F}{N}, \tag{B.2.13}$$

in this form the depencency on the variance of the estimator Var_F is explicit

$$\mathrm{Var}_F = \sum_{\alpha\beta} f_\alpha f_\beta \Gamma_{\alpha\beta}(0)\,, \tag{B.2.14}$$

and multiplied by the integrated autocorrelation

$$A_{I,F} = \frac{1}{2\mathrm{Var}_F} \sum_{t=-\infty}^{\infty} \sum_{\alpha\beta} f_\alpha f_\beta \Gamma_{\alpha\beta}(t). \tag{B.2.15}$$

In absence of autocorrelation we have $\Gamma_{\alpha\beta}(t \neq 0) = 0$ so $A_{I,F} = 1/2$.

B.3 Error estimators

A possible way the extract the error from come correlated data is to build an estimator for $\Gamma_{\alpha\beta}(t)$. To this purpose we start computing

$$\Gamma_{\alpha\beta}(t) = \langle \Gamma_{\alpha\beta}(t) \rangle = \frac{1}{N - Rt} \sum_{r=1}^{R} \sum_{i=1}^{N_r - t} \langle \Gamma_{\alpha\beta}(t) \rangle \tag{B.3.1}$$

$$= \frac{1}{N - Rt} \sum_{r=1}^{R} \sum_{i=1}^{N_r - t} \left\langle (x_\alpha^{i,r} - X_\alpha)(x_\beta^{i+t,r} - X_\beta) \right\rangle = \tag{B.3.2}$$

$$= \frac{1}{N - Rt} \sum_{r=1}^{R} \sum_{i=1}^{N_r - t} \left\langle (x_\alpha^{i,r} - X_\alpha + \bar{\bar{x}}_\alpha - \bar{\bar{x}}_\alpha)(x_\beta^{i+t,r} - X_\beta + \bar{\bar{x}}_\beta - \bar{\bar{x}}_\beta) \right\rangle = \tag{B.3.3}$$

$$= \frac{1}{N - Rt} \sum_{r=1}^{R} \sum_{i=1}^{N_r - t} \Big[\left\langle (x_\alpha^{i,r} - \bar{\bar{x}}_\alpha)(x_\beta^{i+t,r} - \bar{\bar{x}}_\beta) \right\rangle + \langle (\bar{\bar{x}}_\alpha - X_\alpha)(\bar{\bar{x}}_\beta - X_\beta) \rangle + $$
$$+ \left\langle (x_\alpha^{i,r} - \bar{\bar{x}}_\alpha)(\bar{\bar{x}}_\beta - X_\beta) \right\rangle + \left\langle (\bar{\bar{x}}_\alpha - X_\alpha)(x_\beta^{i+t,r} - \bar{\bar{x}}_\beta) \right\rangle \Big] \tag{B.3.4}$$

where in the first line we use the fact that $\Gamma_{\alpha\beta}$ is already an expectation value (B.1.5) and it does not depend on the configuration i and the replica r. As in [53] we can neglect the "end-effect" in the summation of the two term in the last line of the (B.3.4) i.e.

$$\sum_{r=1}^{R} \sum_{i=1}^{N_r - t} (x_\alpha^{i,r} - \bar{\bar{x}}_\alpha) \sim \sum_{r=1}^{R} \sum_{i=1}^{N_r} (x_\alpha^{i,r} - \bar{\bar{x}}_\alpha) = 0. \tag{B.3.5}$$

Defining

$$\bar{\bar{\Gamma}}_{\alpha\beta} = \frac{1}{N - Rt} \sum_{r=1}^{R} \sum_{i=1}^{N_r - t} (x_\alpha^{i,r} - \bar{\bar{x}}_\alpha)(x_\beta^{i+t,r} - \bar{\bar{x}}_\beta), \tag{B.3.6}$$

and using the (B.1.11) and the (B.3.5) the (B.3.4) becomes

$$\Gamma_{\alpha\beta}(t) \sim \left\langle \bar{\bar{\Gamma}}_{\alpha\beta} \right\rangle + \frac{1}{N} C_{\alpha\beta}, \tag{B.3.7}$$

which tells us that $\bar{\bar{\Gamma}}_{\alpha\beta}$ is an estimator of Γ with leading bias $-C_{\alpha\beta}/N$. To cancel this leading order bias we need an estimator for $C_{\alpha\beta}$, we first notice that truncating the sum in $C_{\alpha\beta}$ to a value W

$$C_{\alpha\beta}(W) = \sum_{t=-W}^{W} \Gamma_{\alpha\beta}(t) = \Gamma_{\alpha\beta}(0) + 2 \sum_{t=1}^{W} \Gamma_{\alpha\beta}(t) \tag{B.3.8}$$

we introduce an error that decreases as

$$\frac{C_{\alpha\beta}(W) - C_{\alpha\beta}}{C_{\alpha\beta}} \sim -e^{-W/\tau}. \tag{B.3.9}$$

In general this is an order of magnitude statement but in the case of $\Gamma = e^{-t/\tau}$ can computed directly

$$\frac{C_{\alpha\beta}(W) - C_{\alpha\beta}}{C_{\alpha\beta}} = -\frac{e^{-1/\tau}}{1 + e^{-1/\tau}} e^{-W/\tau} \tag{B.3.10}$$

with the function $e^{-W/\tau}/(1 + e^{-W/\tau})$ is bounded to $(0, 1/2)$. If we chose W large enough we can approximate $C_{\alpha\beta}$ with $C_{\alpha\beta}(W)$, however when W is too large the sum will contain terms with negligible signal and big noise.
An estimator for $C_{\alpha\beta}(W)$ can be

$$\bar{\bar{C}}_{\alpha\beta}(W) = \bar{\bar{\Gamma}}_{\alpha\beta}(0) + 2 \sum_{t=1}^{W} \bar{\bar{\Gamma}}_{\alpha\beta}(t). \tag{B.3.11}$$

Using the (B.3.7) we can compute the leading bias

$$\left\langle \bar{\bar{C}}_{\alpha\beta}(W) \right\rangle - C_{\alpha\beta}(W) \approx -\left(\frac{1 + 2W}{N} \right) C_{\alpha\beta}. \tag{B.3.12}$$

For a derived quantity F we can estimate the projected quantities as

$$\bar{\bar{\Gamma}}_F(t) = \sum_{\alpha\beta} \bar{\bar{f}}_\alpha \bar{\bar{f}}_\beta \bar{\bar{\Gamma}}_{\alpha\beta}(t) \tag{B.3.13}$$

$$\bar{\bar{C}}_F(W) = \Gamma_F(0) + \sum_{t=1}^{W} \bar{\bar{\Gamma}}_F(t) \tag{B.3.14}$$

$$\bar{\bar{A}}_{I,F}(W) = \frac{1}{2\bar{\bar{\Gamma}}_F(0)} \bar{\bar{C}}_F(W) \tag{B.3.15}$$

The error Δ_F^2 (B.2.10) will be given by

$$\bar{\bar{\Delta}}_F^2 = \frac{\bar{\bar{C}}_F(W)}{N} \tag{B.3.16}$$

where $\bar{\bar{f}}_\alpha$ is the derivative of the function f respect X_α computed at $\bar{\bar{x}}_1, \bar{x}_2, \bar{x}_3, ...$, this will introduce an error of $O\left(1/N\right)$. After a choice of W the leading bias in there derived quantities can be eliminated replacing $\bar{\bar{\Gamma}}_{\alpha\beta} \to \mathcal{G}_{\alpha\beta}$.

In some case it may be difficult to compute analytical the dervative of F, a numerical evaluation of it can be done defining

$$h_\alpha = \sqrt{\frac{\bar{\bar{\Gamma}}_{\alpha\alpha}(0)}{4N}} \tag{B.3.17}$$

and taking as estimator with errors $O(h_\alpha^2)$

$$\bar{\bar{f}}_\alpha \approx \frac{1}{2h_\alpha}[f(\bar{\bar{x}}_1, ..., \bar{\bar{x}}_\alpha + h_\alpha, ...) - f(\bar{\bar{x}}_1, ..., \bar{\bar{x}}_\alpha - h_\alpha, ...)]. \tag{B.3.18}$$

B.4 Windowing procedure and leading bias correction

To define a procedure to choose a value for W we first compute the statistical error of the estimator $\bar{\bar{C}}_{\alpha\beta}(W)$

$$\left\langle \bar{\bar{C}}_{\alpha\beta}(W) - C_{\alpha\beta}(W))^2 \right\rangle \approx \frac{2(2W+1)}{N} C^2_{\alpha\beta} \tag{B.4.1}$$

A derivation of this the formula can be find in [54]. This formula was obtained under the main approximation of the four-point function to its disconnected part, which it is reasonable if N is large enough. With the error propagation it is possible to compute also the statistical error in $\bar{\bar{A}}_{I,F}$ (B.3.15)

$$\left\langle (\bar{\bar{A}}_{I,F}(W) - A_{I,F}(W))^2 \right\rangle \approx \frac{4}{N}(W + 1/2 - A_{I,F})A^2_{I,F}. \tag{B.4.2}$$

The statistical error associate with $\Delta^2_{\bar{\bar{X}}_\alpha} = C_{\alpha\alpha}/N$ will be obtained propagating the error on (B.4.1)

$$\delta\Delta^{stat.}_{\bar{\bar{X}}_\alpha} = \frac{1}{2}\sqrt{\frac{2(2W+1)C_{\alpha\alpha}}{N}}. \tag{B.4.3}$$

Besides the statistical error there is the systematic error coming from the truncation in W. From the (B.3.9) we find

$$\delta\Delta^{sys.}_{\bar{\bar{X}}_\alpha} \approx \frac{1}{2} e^{-W/\tau}\sqrt{\frac{C_{\alpha\alpha}}{N}} \tag{B.4.4}$$

The optimal W can be defined as in [3] as the value that minimize the sum of the two error (B.4.3) and (B.4.4)[1]

$$\frac{\delta\Delta^{sys.}_{\bar{\bar{X}}_\alpha} + \delta\Delta^{stat.}_{\bar{\bar{X}}_\alpha}}{\Delta_{\bar{\bar{X}}_\alpha}} \approx \frac{1}{2}\left(e^{-W/\tau} + \sqrt{\frac{2(2W+1)}{N}}\right). \tag{B.4.5}$$

[1] This function is slighty different from the (43) of [3], this difference does effect significant the choice of W

To minimise it we defined implicitly $\bar{\bar{\tau}}(W)$ an estimator for τ trough the equation

$$2\bar{\bar{A}}_{I,F} = \sum_{t=-\infty}^{\infty} e^{-S|t|/\bar{\bar{\tau}}(W)}. \tag{B.4.6}$$

where S is some real number that can be reasonably choose as $S = 1..2$. Inverting the above equation we get

$$\bar{\bar{\tau}}(W) = S\left[\ln\left(\frac{2A_{I,F}(W)+1}{2A_{I,F}(W)-1}\right)\right]^{-1}. \tag{B.4.7}$$

To find the minimum of the (B.4.5) we compute its derivative for increasing value of W

$$g(W) = -\frac{1}{2\bar{\bar{\tau}}(W)}e^{-W/\bar{\bar{\tau}}(W)} + \frac{1}{\sqrt{2N(W+1)}}. \tag{B.4.8}$$

The first value for which $g(W)$ change sign is taken as optimal value on W. In this procedure if $\bar{\bar{A}}_{I,F}(W) \leq 1/2$ the (B.4.7) can't be evaluated so we just set $\bar{\bar{\tau}}$ to a small positive value.

After we find the optima W we can eliminate the leading bias (B.3.7) and (B.3.12) defining

$$\mathcal{G}_{\alpha\beta} = \bar{\bar{\Gamma}}_{\alpha\beta} + \frac{1}{N}\bar{\bar{C}}_{\alpha\beta}(W) \tag{B.4.9}$$

which is an estimator of $\Gamma_{\alpha\beta}$ with bias of order $O\left(1/N^2\right)$

$$\langle\mathcal{G}_{\alpha\beta}\rangle - \Gamma_{\alpha\beta} = \left\langle\bar{\bar{\Gamma}}_{\alpha\beta}\right\rangle - \frac{1}{N}\left\langle\bar{\bar{C}}_{\alpha\beta}(W)\right\rangle - \Gamma_{\alpha\beta} = \tag{B.4.10}$$

$$= \frac{1}{N}C_{\alpha\beta}(W) - \frac{1}{N}\left\langle\bar{\bar{C}}_{\alpha\beta}(W)\right\rangle = O\left(\frac{1}{N^2}\right), \tag{B.4.11}$$

where from the first to the second line we used the (B.3.7) and we approximate $C_{\alpha\beta}$ with $C_{\alpha\beta}(W)$. Similar we can eliminate the leading order bias from in the estimation of $C_{\alpha\beta}(W)$ defining

$$\mathcal{C}_{\alpha\beta}(W) = \mathcal{G}_{\alpha\beta}(0) + 2\sum_{t=1}^{W}\mathcal{G}_{\alpha\beta}(t) = \bar{\bar{C}}_{\alpha\beta}(W)\left(1 + \frac{1+2W}{N}\right). \tag{B.4.12}$$

For the derived quantities we have the

$$\mathcal{G}_F(t) = \sum_{\alpha\beta} \bar{\bar{f}}_\alpha \bar{\bar{f}}_\beta \mathcal{G}_{\alpha\beta}(t) \qquad \langle \mathcal{G}_F(t) \rangle \approx \Gamma_F \tag{B.4.13}$$

$$\mathcal{C}_F(W) = \Gamma_F(0) + \sum_{t=1}^{W} \bar{\bar{\Gamma}}_F(t) \qquad \langle \mathcal{C}_F(W) \rangle \approx C_F \tag{B.4.14}$$

$$\mathcal{A}_{I,F}(W) = \frac{1}{2\bar{\bar{\Gamma}}_F(0)} \bar{\bar{C}}_F(W) \qquad \langle \mathcal{A}_{I,F}(W) \rangle \approx A_{I,F} \tag{B.4.15}$$

The error Δ_F^2 (B.2.10) will be given by

$$\bar{\bar{\Delta}}_F^2 = \frac{\mathcal{C}_F(W)}{N} \tag{B.4.16}$$

Appendix C

Variance in NSPT: an explicit example

As shown by Martin Lüscher [50] the perturbative coefficients $\mathcal{O}_k$ (3.5.11) computed with ISPT have power divergences in their variance. Here we will give a simple example of an observable that his variance diverges as computed with ISPT while it is finite as computed in LSPT. Consider the simple observable $\varphi_1(x)$ defined through the (3.5.3). It can be evaluated in NSPT with the stochastic estimator X

$$X \equiv \frac{1}{V}\sum_x \phi_1(x). \tag{C.0.1}$$

Averaging X over the noise we get

$$\langle X \rangle_\eta = \langle \varphi_1(x) \rangle = 1. \tag{C.0.2}$$

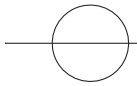
The variance of the estimator X is

$$\mathrm{Var}(X) = \frac{1}{V^2}\sum_{x,y} \left[\langle \phi(x)\phi(y) \rangle_\eta - \langle \phi(x) \rangle_\eta \langle \phi(y) \rangle_\eta \right]. \tag{C.0.3}$$

Since the expectation value of X is zero the variance take the simpler form

$$\mathrm{Var}(X) = \frac{1}{V^2} \sum_{x,y} \left[\langle \phi(x)\phi(y) \rangle_\eta \right]. \tag{C.0.4}$$

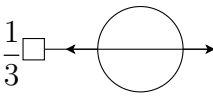
If we calculate the variance of X in ISPT using the diagrammatic expansion in figure 5.1 we get in lattice units

$$\mathrm{Var}(X)^{ISPT} = \frac{1}{24}$$  $$= \frac{1}{24V} \frac{1}{m^4} \sum_{x \in \Omega} \frac{1}{V^3} \left(\sum_{q \in \tilde{\Omega}} \frac{e^{iq \cdot x}}{\hat{q}^2 + m^2} \right)^3, \tag{C.0.5}$$

where the mass renormalization condition (2.3.9) eliminates all the tadpoles. The sunset diagram (C.0.5) is quadratic divergent in the lattice spacing

$$\mathrm{Var}(X)^{ISPT} \propto \frac{1}{a^2}. \tag{C.0.6}$$

Different is the case of LSPT where formally the variance of X in the continuum can be written as a Feynman diagram of the $d+1$ dynamical theory (3.2.1)

$$\mathrm{Var}(X)^{LSPT} = \frac{1}{3}$$  $$= \frac{1}{3} \frac{1}{m^4} \int_{\omega_1,\omega_2} \int_{p_1,p_2} G(\omega_1, p_1) G(\omega_2, p_2) G(\omega_1+\omega_2, p_1+p_2)\,. \tag{C.0.7}$$

Using the power counting rules defined in (3.3.7) we get that the degree of divergent of $\mathrm{Var}(X)^{LSPT}$ is zero, so it is logarithmic divergent if it is computed with the lattice regularization.

Bibliography

[1] M. Dalla Brida, M. Garofalo, and A. D. Kennedy, *Numerical Stochastic Perturbation Theory and Gradient Flow in ϕ^4 Theory*, *PoS* **LATTICE2015** (2015) 309.

[2] M. Dalla Brida, M. Garofalo, and A. D. Kennedy, *Investigation of New Methods for Numerical Stochastic Perturbation Theory in φ^4 Theory*, *Phys. Rev.* **D96** (2017), no. 5 054502, [arXiv:1703.04406].

[3] **ALPHA** Collaboration, U. Wolff, *Monte Carlo errors with less errors*, *Comput. Phys. Commun.* **156** (2004) 143–153. [Erratum: Comput. Phys. Commun. 176 (2007) 383].

[4] I. Montvay and G. Munster, *Quantum fields on a lattice*. Cambridge University Press, 1997.

[5] S. Capitani, *Lattice perturbation theory*, *Phys. Rept.* **382** (2003) 113–302.

[6] M. Lüscher, R. Narayanan, P. Weisz, and U. Wolff, *The Schrödinger functional a renormalizable probe for non-abelian gauge theories*, *Nuclear Physics B* **384** (1992), no. 1 168 – 228.

[7] S. Sint, *On the Schrödinger functional in QCD*, *Nuclear Physics B* **421** (1994), no. 1 135 – 156.

[8] F. Di Renzo, G. Marchesini, P. Marenzoni, and E. Onofri, *Lattice perturbation theory on the computer*, *Nucl. Phys. Proc. Suppl.* **34** (1994) 795–798.

[9] F. Di Renzo, E. Onofri, G. Marchesini, and P. Marenzoni, *Four loop result in SU(3) lattice gauge theory by a stochastic method: lattice correction to the condensate*, *Nucl. Phys.* **B426** (1994) 675–685.

[10] F. Di Renzo and L. Scorzato, *Numerical stochastic perturbation theory for full QCD*, *JHEP* **0410** (2004) 073.

[11] M. Brambilla, M. Dalla Brida, D. Hesse, F. Di Renzo, and S. Sint, *Numerical Stochastic Perturbation Theory for Schrödinger Functional schemes*, *PoS* **LATTICE2013** (2013) 325.

[12] M. Dalla Brida and D. Hesse, *Numerical Stochastic Perturbation Theory and the Gradient Flow*, *PoS* **LATTICE2013** (2014) 326.

[13] G. S. Bali, C. Bauer, A. Pineda, and C. Torrero, *Perturbative expansion of the energy of static sources at large orders in four-dimensional SU(3) gauge theory*, *Phys. Rev.* **D87** (2013) 094517.

[14] G. Parisi and Y. Wu, *Perturbation Theory Without Gauge Fixing*, *Sci. Sin.* **24** (1981) 483.

[15] P. H. Damgaard and H. Hüffel, *Stochastic quantization*, *Physics Reports* **152** (1987), no. 5 227 – 398.

[16] M. Lüscher, *Instantaneous stochastic perturbation theory*, *JHEP* **04** (2015) 142.

[17] P. Weisz and U. Wolff, *Triviality of ϕ_4^4 theory: small volume expansion and new data*, *Nucl. Phys.* **B846** (2011) 316–337.

[18] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory*. Westview, Boulder, CO, 1995. Includes exercises.

[19] M. Lüscher, *Properties and uses of the Wilson flow in lattice QCD*, *JHEP* **08** (2010) 071. [Erratum: JHEP 03 (2014) 092].

[20] M. Lüscher and P. Weisz, *Perturbative analysis of the gradient flow in non-abelian gauge theories*, *JHEP* **02** (2011) 051.

[21] C. Monahan and K. Orginos, *Locally smeared operator product expansions in scalar field theory*, *Phys. Rev.* **D91** (2015), no. 7 074513.

[22] C. Monahan, *The gradient flow in simple field theories*, *PoS* **LATTICE2015** (2015) 052.

[23] Z. Fodor, K. Holland, J. Kuti, D. Nogradi, and C. H. Wong, *The Yang-Mills gradient flow in finite volume*, *JHEP* **11** (2012) 007.

[24] J. Collins, *Renormalization: An Introduction to Renormalization, the Renormalization Group and the Operator-Product Expansion*. Cambridge University Press, 1984.

[25] M. Namiki, *stochastic Quantization.* Springer-Verlag., Berlin Heidelberg, 1992.

[26] M. Lüscher and S. Schaefer, *Non-renormalizability of the HMC algorithm*, *JHEP* **04** (2011) 104.

[27] J. Zinn-Justin, *Renormalization and Stochastic Quantization*, *Nucl. Phys.* **B275** (1986) 135.

[28] A. D. Kennedy, *A Simple proof of the BPH theorem*, in *High-energy physics and quantum field theory. Proceedings, 11th International Workshop, St. Petersburg, Russia, September 12-18, 1996*, pp. 331–338, 1996. `hep-th/9612113`.

[29] W. E. Caswell and A. D. Kennedy, *A simple approach to renormalization theory*, *Phys. Rev.* **D25** (1982) 392.

[30] F. D. Renzo, E. Onofri, G. Marchesini, and P. Marenzoni, *Four-loop result in su(3) lattice gauge theory by a stochastic method: lattice correction to the condensate*, *Nuclear Physics B* **426** (1994), no. 3 675 – 685.

[31] A. S. Kronfeld, *Dynamics of Langevin simulations*, *Prog. Theor. Phys. Suppl.* **111** (1993) 293–312.

[32] G. Aarts and F. A. James, *Complex Langevin dynamics in the SU(3) spin model at nonzero chemical potential revisited*, *JHEP* **01** (2012) 118.

[33] M. Lüscher, *Statistical errors in stochastic perturbation theory*, notes available at luscher.web.cern.ch/luscher/notes/enspt.pdf (2015).

[34] A. M. Horowitz, *Stochastic Quantization in Phase Space*, *Phys. Lett.* **B156** (1985) 89.

[35] S. Duane, *Stochastic quantization versus the microcanonical ensemble: Getting the best of both worlds*, *Nuclear Physics B* **257** (1985) 652 – 662.

[36] A. M. Horowitz, *The Second Order Langevin Equation and Numerical Simulations*, *Nucl. Phys.* **B280** (1987) 510.

[37] A. M. Horowitz, *A Generalized guided Monte Carlo algorithm*, *Phys. Lett.* **B268** (1991) 247–252.

[38] K. Jansen and C. Liu, *Kramers equation algorithm for simulations of QCD with two flavors of Wilson fermions and gauge group SU(2)*, *Nucl. Phys.* **B453** (1995) 375–394. [Erratum: Nucl. Phys. B459 (1996) 437].

[39] M. Lüscher and S. Schaefer, *Lattice QCD without topology barriers*, *JHEP* **07** (2011) 036.

[40] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, *Hybrid Monte Carlo*, *Phys. Lett.* **B195** (1987) 216–222.

[41] A. D. Kennedy and B. Pendleton, *Cost of the Generalized Hybrid Monte Carlo algorithm for free field theory*, *Nucl. Phys.* **B607** (2001) 456–510.

[42] M. A. Clark and A. D. Kennedy, *Asymptotics of Fixed Point Distributions for Inexact Monte Carlo Algorithms*, *Phys. Rev.* **D76** (2007) 074508.

[43] P. B. Mackenzie, *An Improved Hybrid Monte Carlo Method*, *Phys. Lett.* **B226** (1989) 369.

[44] A. D. Kennedy, P. J. Silva, and M. A. Clark, *Shadow Hamiltonians, Poisson Brackets, and Gauge Theories*, *Phys. Rev.* **D87** (2013), no. 3 034511.

[45] I. P. Omelyan, I. M. Mryglod, and R. Folk, *Symplectic analytically integrable decomposition algorithms: classification, derivation, and application to molecular dynamics, quantum and celestial mechanics simulations*, *Comp. Phys. Commun.* **151** (2003) 272.

[46] M. Dalla Brida and M. Lüscher, *SMD-based numerical stochastic perturbation theory*, *Eur. Phys. J.* **C77** (2017), no. 5 308, [arXiv:1703.04396].

[47] G. G. Batrouni, G. R. Katz, A. S. Kronfeld, G. P. Lepage, B. Svetitsky, and K. G. Wilson, *Langevin Simulations of Lattice Field Theories*, *Phys. Rev.* **D32** (1985) 2736.

[48] M. Lüscher, *luscher.web.cern.ch/luscher/ISPT/ISPT-0.1.tar.gz*, .

[49] M. Frigo and S. G. Johnson, *The Design and Implementation of FFTW3*, *Proceedings of the IEEE* **93** (2005), no. 2 216–231. Special issue on "Program Generation, Optimization, and Platform Adaptation".

[50] M. Lüscher, *Numerical stochastic perturbation theory revisited*, Talk given at: *Fundamental parameters from lattice QCD,* Mainz (2015).

[51] M. Dalla Brida and M. Lüscher, *The gradient flow coupling from numerical stochastic perturbation theory*, *PoS* **LATTICE2016** (2016) 332.

[52] R. Frezzotti, M. Hasenbusch, U. Wolff, J. Heitger, and K. Jansen, *Comparative benchmarks of full qcd algorithms*, *Computer Physics Communications* **136** (2001), no. 1 1 – 13.

[53] M. B. Priestley, *Spectral Analysis and Time Series.* Academic Press INC., London, 1981.

[54] N. Madras and A. D. Sokal, *The Pivot algorithm: a highly efficient Monte Carlo method for selfavoiding walk*, *J. Statist. Phys.* **50** (1988) 109–186.